



# Optimisation, analyse et comparaison de méthodes numériques déterministes par la dynamique des gaz raréfiés

Nicolas Herouard

## ► To cite this version:

Nicolas Herouard. Optimisation, analyse et comparaison de méthodes numériques déterministes par la dynamique des gaz raréfiés. Mathématiques générales [math.GM]. Université de Bordeaux, 2014. Français. NNT : 2014BORD0473 . tel-01163271

**HAL Id: tel-01163271**

**<https://theses.hal.science/tel-01163271>**

Submitted on 12 Jun 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE

présentée à

**L'UNIVERSITÉ DE BORDEAUX**

**École doctorale de Mathématiques et Informatique**

par

**Nicolas HÉROUARD**

pour l'obtention du titre de

**DOCTEUR**

**Spécialité : Mathématiques Appliquées et Calcul Scientifique**

---

**Optimisation, analyse et comparaison  
de méthodes numériques déterministes  
pour la dynamique des gaz raréfiés**

---

Soutenue le 5 décembre 2014

devant le jury composé de :

Luc MIEUSSENS	Professeur, Université de Bordeaux	Directeur de thèse
Laurent DESVILLETES	Professeur, ENS Cachan	Rapporteur
Thierry MAGIN	Professeur, Institut Von Kármán	Rapporteur
Céline BARANGER	Ingénieur de recherche, CEA-CESTA	Encadrante
Julien MATHIAUD	Ingénieur de recherche, CEA-CESTA	Encadrant
Rodolphe TURPAULT	Professeur, Université de Bordeaux	Président
Stéphane DELLACHERIE	Ingénieur de recherche, CEA Saclay	Examineur
Florian MÉHATS	Professeur, Université de Rennes 1	Examineur



# Optimisation, analyse et comparaison de méthodes numériques déterministes pour la dynamique des gaz raréfiés

## Résumé

Lors de la rentrée atmosphérique, l'écoulement raréfié de l'air autour de l'objet rentrant est régi par un modèle cinétique dérivé de l'équation de Boltzmann ; celui-ci décrit l'évolution d'une fonction de distribution des particules de gaz dans l'espace des phases, de dimension 6 dans le cas général. La simulation numérique déterministe de cet écoulement requiert donc le traitement d'une quantité considérable de données, soit un espace mémoire et un temps de calcul importants. Nous étudions dans ce travail différents moyens de réduire le coût de ces calculs. La première approche est une méthode permettant d'optimiser la taille de la grille de vitesses discrètes employée dans le calcul par une prédiction de l'allure des fonctions de distribution dans l'espace des vitesses, en supposant un faible déséquilibre thermodynamique du gaz. La seconde approche consiste à essayer d'exploiter les propriétés de préservation asymptotique des schémas Galerkin Discontinu, déjà établies dans le cadre du transport linéaire des neutrons, qui permettent de tenir compte des effets de la couche limite cinétique sans que celle-ci soit résolue par le maillage, alors que les méthodes classiques (comme les Volumes Finis) imposent l'utilisation de maillages très raffinés en zone de proche paroi. Dans une dernière partie, nous comparons les performances respectives de ces schémas Galerkin Discontinu et de quelques schémas Volumes Finis, appliqués au modèle BGK sur un cas simple, en étudiant en particulier leur comportement près des parois et les conditions aux limites numériques.

**Mots-clés :** aérodynamique, régime raréfié, méthodes numériques, schémas Galerkin Discontinu, limite asymptotique, coût de calcul, équation de Boltzmann-BGK, conditions aux limites

# Optimization, analysis and comparison of deterministic numerical methods for rarefied gas dynamics

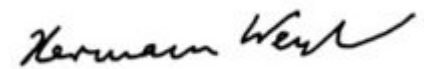
## Abstract

During the atmospheric re-entry of a space engine, the rarefied air flow around the body is determined by a kinetic model derived from the Boltzmann equation, which describes the evolution of a distribution function of gas molecules in the phase space, this means a 6-dimensional space in the general case. Consequently, a deterministic numerical simulation of this flow requires large computational resources, both in memory storage and CPU time. The aim of this work is to reduce those resources, using two different approaches. The first one is a method allowing to optimize the size of the discrete velocity grid used for the computation by a prediction of the shape of the distributions in the velocity space, assuming that the gas is close to thermodynamic equilibrium. The second approach is an attempt to use the asymptotic preservation properties of Discontinuous Galerkin schemes, already established for neutron transport, which allow to take into account the effects of kinetic boundary layers even if they are not resolved by the mesh, while classical methods (such as Finite Volumes) require very refined meshes along the direction normal to the walls. In a last part, we compare the performances of these Discontinuous Galerkin schemes with some classical Finite Volumes schemes, applied to the BGK equation in a simple case, and pay particular attention to their near-wall behavior and numerical boundary conditions.

**Keywords :** aerodynamics, rarefied regime, Boltzmann-BGK equation, numerical methods, Discontinuous Galerkin schemes, asymptotic limit, computational cost, boundary conditions



*Mon travail a toujours cherché à concilier vérité et beauté,  
mais lorsque j'avais à choisir entre l'une et l'autre,  
généralement je préférais la beauté.*

A handwritten signature in black ink, reading "Hermann Weyl". The script is fluid and cursive, with a prominent flourish at the end of the word "Weyl".

Hermann Weyl



# Remerciements

Arrivé au terme de ce travail, le temps est venu de remercier tous ceux qui ont contribué à son élaboration et à son achèvement, de près ou de loin.

À tout seigneur tout honneur, mes premiers remerciements vont à Luc Mieussens, qui a dirigé mon travail de thèse durant ces trois années et demie avec rigueur, patience, mais aussi et surtout avec gentillesse et humour.

Je suis très reconnaissant à Laurent Desvilletes et à Thierry Magin d'avoir accepté de rapporter cette thèse au moment de la rentrée (universitaire et non pas atmosphérique celle-là). Merci aussi aux examinateurs d'avoir bien voulu faire partie du jury.

Je remercie également les chefs de laboratoire et de service de m'avoir accueilli au CEA-CESTA, en particulier Pierre-Henri Maire qui m'a orienté (après quelques péripéties) sur le stage de fin d'études qui a débouché sur cette thèse et m'a ainsi permis de découvrir le petit monde de la dynamique des gaz raréfiés, qui était alors pour moi une *terra incognita* que j'ai (modestement) explorée avec les yeux d'un Christophe Colomb débarqué dans le Nouveau Monde. Je n'oublie bien entendu pas la gentillesse de Xavier Carlotti.

Je remercie naturellement Céline Baranger qui a suivi mon travail au CESTA. Sa disponibilité, ses conseils, ses encouragements et sa bonne humeur m'ont été précieux.

Cette thèse ne serait rien sans l'aide inestimable de Julien Mathiaud, et les longues heures passées ensemble (sans s'endormir) à déboguer du code, à noircir le tableau de mon bureau de schémas et de calculs qui partent dans tous les sens, à s'interroger sur la moralité des résultats. La ponctuation de ce manuscrit lui doit également beaucoup... Et le fouet n'a pas tant servi que ça, finalement !

Je remercie bien sûr toutes les autres personnes que j'ai côtoyées au sein du service pendant ces années, aussi bien les chercheurs permanents (mention particulière à Thanh-Hà, Marc, Jean et Gérard) que les thésards : ceux qui sont toujours là, Gabriel (le poil à gratter du labo) et Matthieu (le ch'ti qu'on aime bien quand même), et ceux qui sont partis vers de nouvelles aventures, François V. et François G. (l'amateur de littérature russe), Manu (qui a répondu à un appel plus fort que celui de la science), Pascal, Cyril, Thibault, Mathieu... Merci également à Guilhem Caramel, et bien sûr à Anne-Pascale Leroux, secrétaire irremplaçable !

Je n'oublie pas mon *alter ego* de l'IMB, Guillaume Dechristé, à qui mon étude des conditions aux limites sur les schémas Volumes Finis appliqués au modèle BGK à vitesses discrètes doit énormément, et qui a été un agréable compagnon de conférences à Saragosse et à Oxford.

*Last but not least*, un grand merci à mes parents qui m'ont toujours soutenu et encouragé dans mes choix et mes projets au cours de mes études. Cette thèse, qui vient les clôturer, n'aurait jamais vu le jour sans cela.





# Table des matières

<b>Remerciements</b>	<b>vi</b>
<b>Table des matières</b>	<b>viii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Raffinement local des grilles de vitesses discrètes</b>	<b>13</b>
1.1 Introduction . . . . .	13
1.2 Boltzmann equation and Cartesian velocity grid . . . . .	15
1.2.1 Kinetic description of rarefied gases . . . . .	15
1.2.2 Velocity discretization on Cartesian grid . . . . .	16
1.3 An algorithm to define locally refined discrete velocity grids . . . . .	17
1.3.1 The support function . . . . .	18
1.3.2 AMR grid generation . . . . .	19
1.3.3 An example . . . . .	20
1.3.4 Discrete Velocity Approximation on the AMR velocity grid . . . . .	20
1.3.5 AMR grid generation for axisymmetric flows . . . . .	21
1.4 Numerical results . . . . .	22
1.4.1 The 3D kinetic code . . . . .	22
1.4.2 A 2D plane example . . . . .	23
1.4.3 A 2D axisymmetric example . . . . .	25
1.4.4 A 3D example . . . . .	26
1.4.5 Initialization with the CNS results . . . . .	26
1.5 Conclusion and perspectives . . . . .	27
<b>2 Analyse asymptotique de l'équation BGK linéarisée avec condition de réflexion diffuse aux parois</b>	<b>39</b>
2.1 Adimensionnement de l'équation BGK stationnaire . . . . .	40
2.2 Linéarisation de l'équation adimensionnée . . . . .	43
2.2.1 Expression formelle : perturbation des fonctions de distribution . . . . .	43
2.2.2 Expression de l'opérateur BGK linéarisé en fonction des moments . . . . .	44
2.3 Limite incompressible . . . . .	47
2.4 Prise en compte des conditions aux limites . . . . .	50
2.4.1 Expression adimensionnelle des conditions aux limites aux parois . . . . .	50
2.4.2 Conditions aux limites et limite incompressible : couche limite cinétique . . . . .	51
2.4.2.1 Problème BGK linéarisé avec conditions aux limites . . . . .	51

2.4.2.2	Développement de Grad-Hilbert : notion de couche limite cinétique	52
2.4.3	Correcteur de couche limite . . . . .	54
2.4.3.1	Expression du correcteur . . . . .	54
2.4.3.2	Variables de couche limite et problèmes de demi-espace . . . . .	55
2.4.3.3	Correcteur au premier ordre : conditions de glissement . . . . .	58
2.4.3.4	Glissement thermique : discussion du signe de la constante $K_3$ . . . . .	65
2.5	Bilan . . . . .	66
<b>3</b>	<b>Application de la méthode Galerkin Discontinu au problème BGK linéarisé avec condition aux limites de réflexion diffuse</b>	<b>69</b>
3.1	Rappel du problème aux limites linéarisé . . . . .	70
3.2	Formulation variationnelle du jeu d'équations de Stokes avec conditions aux limites	72
3.2.1	Problème de Laplace : champ $T'_0$ . . . . .	72
3.2.2	Problème de Boussinesq : champ $\rho'_0$ . . . . .	73
3.2.3	Problème de Stokes sous contrainte : champ $\mathbf{u}_0'$ . . . . .	74
3.3	Écriture du schéma Galerkin Discontinu . . . . .	77
3.3.1	Forme faible de l'équation BGK linéarisée . . . . .	78
3.3.2	Approximation Galerkin Discontinu . . . . .	79
3.3.2.1	Définitions et notations . . . . .	79
3.3.2.2	Construction du schéma Galerkin Discontinu . . . . .	80
3.4	Analyse asymptotique du schéma Galerkin Discontinu . . . . .	83
3.4.1	Développement de Hilbert . . . . .	83
3.4.2	Termes d'ordre -1 . . . . .	85
3.4.3	Termes d'ordre 0 . . . . .	85
3.4.3.1	Continuité de $f_0^h$ et conditions aux limites fortes . . . . .	85
3.4.3.2	Remarque sur l'hypothèse d'uniformité de $\mathbf{u}_w$ et $\theta_w$ . . . . .	91
3.4.3.3	Équations de continuité et de Boussinesq . . . . .	92
3.4.4	Termes d'ordre 1 . . . . .	93
3.5	Convergence des problèmes discrets . . . . .	96
3.5.1	Problème de Laplace . . . . .	97
3.5.2	Problème de Boussinesq . . . . .	98
3.5.3	Problème de Stokes . . . . .	98
3.6	Analyse asymptotique : termes d'ordre $n \geq 2$ . . . . .	100
3.7	Bilan . . . . .	101
<b>4</b>	<b>Résolution numérique de l'équation BGK 1D : écoulement de Couette</b>	<b>103</b>
4.1	Modèle à distributions marginales . . . . .	103
4.1.1	Équation pour la distribution $f$ . . . . .	105
4.1.2	Équation pour la distribution $g$ . . . . .	106
4.1.3	Équation pour la distribution $h$ . . . . .	107
4.1.4	Conditions aux limites . . . . .	108
4.1.5	Système final d'équations . . . . .	109
4.2	Schéma implicite linéarisé . . . . .	110
4.2.1	Écriture et linéarisation . . . . .	110
4.2.2	Expression des différents termes du schéma . . . . .	112
4.2.2.1	Distributions d'équilibre discrètes . . . . .	112
4.2.2.2	Différentielles des distributions d'équilibre . . . . .	113

4.2.2.3	Termes diagonaux de la matrice de relaxation $\bar{\bar{R}}^n$ (matrice $\bar{\bar{\Delta}}^n$ ) .	115
4.3	Schémas Volumes Finis : flux et condition de réflexion diffuse . . . . .	115
4.3.1	Schémas d'ordre 1 . . . . .	116
4.3.2	Schémas d'ordre 2 . . . . .	117
4.3.2.1	Schémas à reconstruction linéaire . . . . .	117
4.3.2.2	Schémas à limiteur de flux . . . . .	121
4.3.2.3	Comparaison des deux types de schémas . . . . .	122
4.3.3	Schéma d'ordre 4 . . . . .	125
4.4	Schéma Galerkin Discontinu . . . . .	126
4.4.1	Ecriture du schéma semi-discret . . . . .	127
4.4.1.1	Terme instationnaire . . . . .	127
4.4.1.2	Terme de flux . . . . .	128
4.4.1.3	Terme de collision . . . . .	129
4.4.1.4	Formulation matricielle . . . . .	130
4.4.2	Discrétisation temporelle . . . . .	132
4.4.3	Conditions aux limites . . . . .	132
4.5	Comparaison des schémas Volumes Finis et Galerkin Discontinu . . . . .	132
4.5.1	Maillages réguliers : estimation de l'erreur et ordres de convergence . . . .	133
4.5.2	Maillages non réguliers . . . . .	140
4.5.3	Vitesse de convergence en temps : analyse des valeurs et espaces propres du système linéaire . . . . .	141
4.6	Conclusion . . . . .	144
<b>Conclusion</b>		<b>145</b>
<b>A Overview of the 3D kinetic code</b>		<b>147</b>
A.1	The linearized implicit scheme . . . . .	147
A.2	Numerical boundary conditions . . . . .	149
A.3	Linear solver . . . . .	151
A.4	Jacobian matrices and the relaxation matrix $R^n$ . . . . .	153
<b>B Équilibre linéarisé et opérateur <math>\Pi</math></b>		<b>155</b>
B.1	Linéarisation de la maxwellienne . . . . .	155
B.2	Moments de l'opérateur $\Pi(\rho', \mathbf{u}', T')$ . . . . .	156
<b>C Modèle à 2 vitesses pour le correcteur de couche limite</b>		<b>159</b>
<b>D Formulaire</b>		<b>165</b>
<b>Bibliographie</b>		<b>167</b>



*À Grégoire Taviot et Pierre Comte<sup>†</sup>,  
avec reconnaissance.*

*À Jorge Colman Lerner<sup>†</sup>,  
avec affection.*



# Introduction

*Il y a un but, mais pas de chemin ; ce que nous nommons chemin est hésitation.*

Franz Kafka - *Préparatifs de noce à la campagne*

Lors de sa rentrée dans l'atmosphère terrestre, un objet spatial traverse différentes couches d'air de caractéristiques physiques très différentes, depuis l'ionosphère composée de gaz ionisés de très faible densité jusqu'aux couches basses de l'atmosphère où l'air est un gaz neutre assimilable à un milieu continu. On utilise en conséquence des modèles physiques différents pour décrire l'écoulement de l'air autour de l'objet rentrant, d'un point de vue dynamique mais également pour tenir compte d'effets chimiques ou thermomécaniques (comme la dissociation des molécules de gaz ou l'ablation de la paroi).

En ce qui concerne l'aspect purement aérodynamique, l'objet en cours de rentrée traverse des couches d'air de densité croissante, avec une variation importante entre le moment de son entrée dans l'atmosphère et l'approche du sol. En conséquence, l'écoulement autour du corps passe par différents régimes, caractérisés par le niveau de raréfaction de l'air. Celui-ci est décrit par un paramètre adimensionnel, le nombre de Knudsen, noté  $\text{Kn}$ , égal au rapport entre le libre parcours moyen  $\lambda$  des molécules d'air, c'est-à-dire la distance moyenne parcourue par ces molécules entre deux collisions avec leurs voisines, et la longueur caractéristique de l'objet rentrant  $L$ , généralement de l'ordre du mètre ou de la dizaine de mètres ([43], d'après une intuition de Reynolds [61]) :

$$\text{Kn} = \frac{\lambda}{L}.$$



FIGURE 1: MARTIN KNUDSEN (1871-1949).

On peut également définir un nombre de Knudsen local en utilisant comme longueur de référence  $L$  la distance caractéristique de variation des grandeurs thermodynamiques au sein de l'écoulement ; on choisit généralement pour cela la longueur caractéristique issue du gradient de la



densité  $\rho$  du gaz :

$$L = \frac{\rho}{\|\nabla \rho\|_{L^2}}.$$

On distingue quatre régimes principaux d'écoulement en fonction de l'ordre de grandeur du nombre de Knudsen :

- $\text{Kn} > 10$  : **régime moléculaire libre**. Les particules de gaz sont libres de se mouvoir sans quasiment jamais entrer en collision les unes avec les autres.
- $1 < \text{Kn} < 10$  : **régime raréfié**. Les molécules se déplacent selon une trajectoire rectiligne entre deux chocs avec les molécules environnantes. Ceux-ci sont trop espacés pour que le gaz soit considéré comme un milieu continu.
- $0.1 < \text{Kn} < 1$  : **régime transitionnel**. Le gaz forme un milieu "quasi-continu", sauf dans les chocs ou les couches limites où le nombre de Knudsen local reste élevé à cause des variations rapides des grandeurs macroscopiques.
- $\text{Kn} < 0.1$  : **régime continu** (ou régime fluide).

À titre illustratif, on montre sur la figure (3) une estimation de la gamme d'altitudes correspondant à ces différents régimes pour une longueur caractéristique de l'objet  $L = 1$  m.

Dans le régime fluide, la dynamique de l'air est parfaitement décrite au niveau macroscopique par les équations d'Euler ou de Navier-Stokes, liant entre elles les grandeurs thermodynamiques (densité  $\rho$ , vitesse  $\mathbf{u}$ , densité d'énergie  $e$ , température  $T$ , pression  $p$ , flux de chaleur  $\mathbf{q}^1$ ) ; à ces équations sont adjointes des lois d'état décrivant la réponse locale du milieu continu à certaines grandeurs (comme la loi des gaz parfaits pour la pression) et des expressions macroscopiques des phénomènes de diffusion (loi de Fourier pour le flux de chaleur), ce qui permet de constituer un système d'équations fermé. Dans les autres régimes, la non-localité impose d'utiliser une description au niveau mésoscopique. On s'intéresse alors à la fonction de distribution  $f$  des molécules, correspondant à la densité de masse de particules dans l'espace des phases ; la grandeur  $f(t, \mathbf{x}, \mathbf{v})d\mathbf{x}d\mathbf{v}$  est ainsi la masse de particules contenue au temps  $t$  dans un volume élémentaire de cet espace autour de la position  $\mathbf{x}$  et de la vitesse  $\mathbf{v}$ . L'évolution de cette fonction de distribution est décrite par l'équation de Boltzmann, fondement de la théorie cinétique des gaz, introduite par Ludwig Boltzmann en 1872 [15].

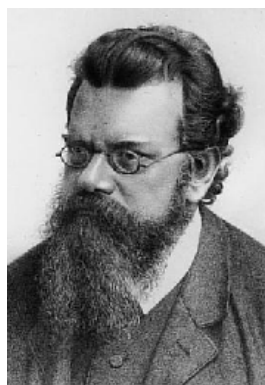


FIGURE 2: LUDWIG BOLTZMANN (1844-1906).

---

1. Dans tout ce document, on adopte les notations suivantes : les quantités scalaires sont notées  $a$ , les quantités vectorielles  $\mathbf{a}$  et les quantités tensorielles  $\bar{\bar{A}}$ .

Cette équation s'écrit en l'absence de champ de force externe (on néglige par exemple la gravité)

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_x f = Q(f), \quad (1)$$

où  $Q(f)$  est un terme de collision traduisant l'effet des chocs entre particules sur l'évolution de la fonction de distribution  $f$ . Dans le régime moléculaire libre, ces collisions étant négligées, on a  $Q(f) = 0$ .

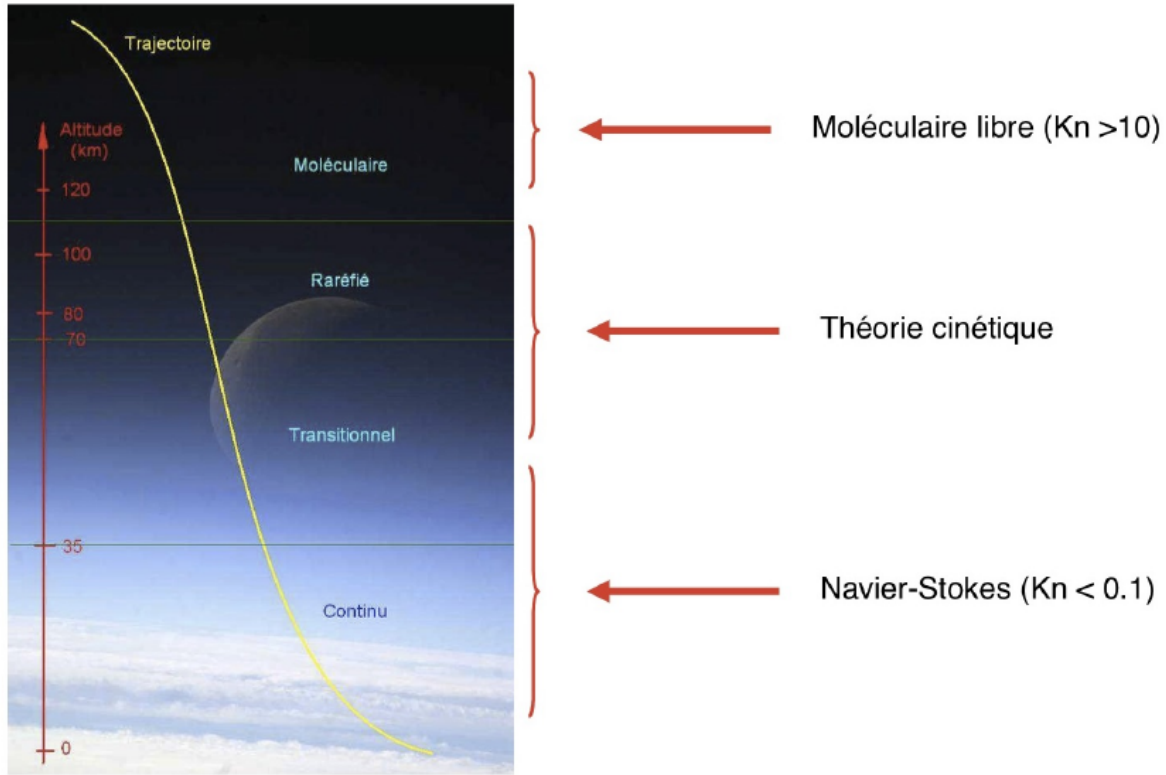


FIGURE 3: Les différents régimes d'écoulement lors de la rentrée atmosphérique d'un objet spatial de longueur caractéristique  $L = 1$  m.

La fonction de distribution  $f$  est reliée aux grandeurs thermodynamiques conservatives (densité, quantité de mouvement, énergie totale) par ses trois premiers moments en fonction de la variable de vitesse  $\mathbf{v}$ . En introduisant la notation

$$\langle g \rangle = \int_{\mathbb{R}^3} g(t, \mathbf{x}, \mathbf{v}) d\mathbf{v},$$

et le vecteur des invariants collisionnels

$$\mathbf{m}(\mathbf{v}) = \left( 1, \mathbf{v}, \frac{1}{2} |\mathbf{v}|^2 \right)^T,$$

on définit le vecteur des quantités conservatives

$$\boldsymbol{\rho}(t, \mathbf{x}) = \begin{pmatrix} \rho(t, \mathbf{x}) \\ \rho \mathbf{u}(t, \mathbf{x}) \\ \rho e(t, \mathbf{x}) \end{pmatrix} = \left\langle \mathbf{m}(\mathbf{v}) f(t, \mathbf{x}, \mathbf{v}) \right\rangle. \quad (2)$$

On définit également le tenseur des contraintes  $\bar{\bar{P}}$  et le vecteur flux de chaleur  $\mathbf{q}$  par

$$\begin{cases} \bar{\bar{P}} = \langle (\mathbf{u} - \mathbf{v}) \otimes (\mathbf{u} - \mathbf{v}) f \rangle, \\ \mathbf{q} = \frac{1}{2} \langle |\mathbf{u} - \mathbf{v}|^2 (\mathbf{u} - \mathbf{v}) f \rangle ; \end{cases} \quad (3)$$

la pression s'exprime alors

$$p = \frac{1}{3} \text{Tr}(\bar{\bar{P}}) = \frac{1}{3} \langle |\mathbf{u} - \mathbf{v}|^2 f \rangle ;$$

celle-ci étant reliée à la température  $T$  par la loi des gaz parfaits, on déduit naturellement

$$T = \frac{p}{\rho R} = \frac{1}{3\rho R} \langle |\mathbf{u} - \mathbf{v}|^2 f \rangle .$$

L'énergie totale obtenue en (2.17) s'exprime donc

$$\rho e = \frac{1}{2} \rho |\mathbf{u}|^2 + \frac{3}{2} \rho RT .$$

On montre que pour toute distribution  $f \geq 0$ , l'opérateur de collision vérifie les trois propriétés suivantes (voir [21])

$$\begin{cases} \langle \mathbf{m} Q(f) \rangle = \mathbf{0}, \\ \langle Q(f) \log f \rangle \leq 0 \quad (\text{théorème H}), \\ (Q(f) = 0) \Leftrightarrow (f = M(f)), \end{cases} \quad (4)$$

où la distribution d'équilibre local  $M(f)$  est une distribution maxwellienne [49], et s'exprime en fonction des grandeurs thermodynamiques associées à la distribution  $f$  (densité, vitesse et température)

$$M(f) = M[\rho, \mathbf{u}, T] = \frac{\rho}{(2\pi RT)^{3/2}} \exp \left( -\frac{|\mathbf{u} - \mathbf{v}|^2}{2RT} \right),$$

avec  $R$  la constante spécifique du gaz, définie à partir de la constante de Boltzmann  $k_B$ , du nombre d'Avogadro  $\mathcal{N}_A$  et de la masse molaire  $\mathcal{M}$  du gaz :  $R = \frac{k_B \mathcal{N}_A}{\mathcal{M}}$ .



FIGURE 4: JAMES CLERK MAXWELL (1831-1879).

La première égalité du système (4) donne, en multipliant l'équation de Boltzmann (1) par  $\mathbf{m}$  avant de l'intégrer par rapport à  $\mathbf{v}$ , les lois de conservation

$$\frac{\partial}{\partial t} \langle \mathbf{m} f \rangle + \nabla \cdot \langle \mathbf{m} \otimes \mathbf{v} f \rangle = \mathbf{0} ; \quad (5)$$

la seconde donne, en définissant l'entropie cinétique par

$$H(f) = \langle f \log f \rangle$$

la loi de décroissance de l'entropie

$$\frac{\partial}{\partial t} H(f) + \nabla \cdot \langle \mathbf{v} f \log f \rangle \leq 0. \quad (6)$$

Les lois de conservation (5) s'écrivent avec les grandeurs thermodynamiques

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \\ \frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u} + \bar{\bar{P}}) = \mathbf{0}, \\ \frac{\partial \rho e}{\partial t} + \nabla \cdot (\rho e \mathbf{u} + \bar{\bar{P}} \mathbf{u} + \mathbf{q}) = 0. \end{cases} \quad (7)$$

A cause des moments du troisième ordre dans l'expression du vecteur  $\mathbf{q}$ , le système n'est pas fermé.

Le terme de collision  $Q(f)$  est très complexe, et pose de nombreuses difficultés tant au niveau théorique que numérique. Différents modèles cinétiques plus simples ont donc été proposés pour remplacer l'équation de Boltzmann : moyennant certaines hypothèses de validité, ces modèles décrivent l'effet des collisions au niveau macroscopique plutôt qu'au niveau mésoscopique, tout en respectant les lois (5) et (6), ainsi que la positivité de  $f$  (mais sans toutes posséder le même équilibre). Pour l'aérodynamique en régime transitionnel, le modèle le plus répandu, auquel on s'intéresse ici, est le modèle BGK, introduit par Bhatnagar, Gross et Krook en 1954 [13] (voir [58] pour la preuve de l'existence d'une solution globale). Ce modèle, parfois appelé dans la littérature modèle de Boltzmann-Krook-Welander, ou BKW, assimile le terme de collision à une relaxation des distributions vers l'équilibre thermodynamique associé  $M(f)$ , avec un temps de relaxation  $\tau$  dépendant des grandeurs thermodynamiques (généralement de la densité et de la température) :

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_x f = \frac{1}{\tau} (M(f) - f).$$

A ces modèles cinétiques sont adjointes différentes conditions aux limites. Dans ce travail, on va s'intéresser tout particulièrement à l'interaction entre le gaz raréfié et les parois de l'engin de rentrée ; on ne décrit donc ici que la condition aux limites traduisant cette interaction.

Les particules de gaz arrivant sur la paroi subissent une réflexion, qui peut s'effectuer selon deux modes, la réflexion spéculaire ou la réflexion diffuse. Dans la réflexion spéculaire, on considère que les particules de gaz rebondissent sur la paroi comme des boules de billard, suivant les lois de Descartes : il n'y a alors aucun échange d'énergie entre le gaz et la paroi  $\Gamma$ . Cette condition s'exprime mathématiquement

$$f(t, \mathbf{x}, \mathbf{v})|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = f(t, \mathbf{x}, R(\mathbf{v})), \quad R(\mathbf{v}) = \mathbf{v} - 2(\mathbf{v} \cdot \mathbf{n})\mathbf{n},$$

où  $\mathbf{n}$  est la normale sortante à la paroi (dirigée de la paroi vers le gaz).

Dans la réflexion diffuse en revanche, les particules de gaz sont thermalisées, et réémises suivant une distribution maxwellienne calculée sur la température et la vitesse de la paroi, notées respectivement  $T_w$  et  $\mathbf{u}_w$ , cette dernière étant supposée tangente à la paroi ( $\mathbf{u}_w \cdot \mathbf{n} = 0$ ) :

$$f(t, \mathbf{x}, \mathbf{v})|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = \sigma(f) M(1, \mathbf{u}_w, T_w),$$

où le coefficient de réflexion  $\sigma(f)$  est égal au rapport des demi-flux de masse entrant et sortant à la paroi, ce qui assure que le flux de masse total de particules de gaz traversant la paroi est nul :

$$\sigma(f) = - \left( \int_{\mathbf{v} \cdot \mathbf{n} < 0} (\mathbf{v} \cdot \mathbf{n}) f d\mathbf{v} \right) \left( \int_{\mathbf{v} \cdot \mathbf{n} > 0} (\mathbf{v} \cdot \mathbf{n}) M(1, \mathbf{u}_w, T_w) d\mathbf{v} \right)^{-1} = - \sqrt{\frac{2\pi}{RT_w}} \int_{\mathbf{v} \cdot \mathbf{n} < 0} (\mathbf{v} \cdot \mathbf{n}) f d\mathbf{v}.$$

La réflexion diffuse entraîne donc, via cette thermalisation, un échange de chaleur entre le gaz et la paroi. Lors d'une rentrée atmosphérique, où l'engin spatial atteint des vitesses très importantes (régime hypersonique), ce flux de chaleur est considérable (à titre d'exemple, on voit sur la figure (5) le bouclier thermique du module de contrôle "Charlie Brown" de la capsule Apollo 10 (1969), carbonisé lors de la rentrée). Il est donc très important pour l'ingénierie de pouvoir calculer précisément ce flux de chaleur afin d'éviter la destruction du corps de rentrée.



FIGURE 5: Module de commande d'Apollo 10 (Science Museum, Londres).

La condition de réflexion générale la plus classique est un couplage de la réflexion spéculaire et de la réflexion diffuse, appelé condition d'accomodation de Maxwell [50] :

$$f(t, \mathbf{x}, \mathbf{v})|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = \alpha f(t, \mathbf{x}, R(\mathbf{v})) + (1 - \alpha) \sigma(f) M(1, \mathbf{u}_w, T_w),$$

où la constante  $\alpha \in [0, 1]$  est le coefficient d'accomodation. La détermination de la valeur de  $\alpha$  pour une configuration donnée est un problème complexe dépendant de la nature et de l'état de la paroi ; dans le cas de la rentrée atmosphérique, où la paroi du corps de rentrée présente de fortes rugosités à cause de l'ablation, la réflexion diffuse est nettement dominante et  $\alpha$  dépasse rarement une valeur de 25%. Cette prédominance de la réflexion diffuse ainsi que son rôle exclusif dans les échanges de chaleur à la paroi du corps de rentrée justifient que l'on néglige dans notre étude la réflexion spéculaire.

Il existe des modèles de réflexion générale plus élaborés, avec un coefficient d'accommodation dépendant de  $\mathbf{v}$  comme celui de Struchtrup [66], ou à plusieurs coefficients comme celui de Cercignani et Lampis [22], mais leur utilisation reste marginale.

Il y a naturellement une compatibilité entre les équations cinétiques utilisées dans le régime transitionnel et les équations de Navier-Stokes utilisées dans le régime fluide, la description cinétique étant également valable dans ce régime. En effectuant un développement asymptotique de Chapman-Enskog [23] sur l'équation de Boltzmann, on montre que les équations de Navier-Stokes compressibles sont une approximation d'ordre 2 de cette équation en fonction du nombre de Knudsen, supposé petit (les équations d'Euler compressibles, négligeant les phénomènes diffusifs et les transferts thermiques, sont une approximation d'ordre 1 et peuvent s'obtenir avec un développement de Hilbert, plus général que celui de Chapman-Enskog). Le tenseur  $\bar{\bar{P}}$  et le vecteur  $\mathbf{q}$  obtenus en (3) s'écrivent dans cette approximation

$$\begin{cases} \bar{\bar{P}} = \rho RT \bar{\bar{I}} - \varepsilon \mu \left( \nabla \mathbf{u} + \nabla \mathbf{u}^T - \frac{2}{3} \nabla \cdot \mathbf{u} \bar{\bar{I}} \right) + \mathcal{O}(\varepsilon^2), \\ \mathbf{q} = -\varepsilon \kappa \nabla T + \mathcal{O}(\varepsilon^2) \quad (\text{loi de Fourier}), \end{cases} \quad \varepsilon = \text{Kn} \ll 1. \quad (8)$$

En revanche, en adjoignant aux équations de Navier-Stokes les conditions aux limites "standard" d'adhérence à la paroi, on n'obtient qu'une approximation d'ordre 1 de l'équation de Boltzmann avec des conditions de réflexion. Lorsque Kn n'est plus négligeable, ces conditions aux limites ne sont donc plus valides. On voit alors apparaître le long de la paroi du corps de rentrée une couche limite cinétique, appelée couche de Knudsen. Une analyse de cette couche limite avec la décomposition de Chapman-Enskog permet d'établir de nouvelles conditions aux limites pour les grandeurs thermodynamiques, caractérisées par une discontinuité entre la vitesse et la température du fluide et celles de la paroi : on les nomme "conditions de glissement". On distingue alors deux sous-régimes au régime fluide :

- $0.01 < \text{Kn} < 0.1$  : **régime de glissement**. L'écoulement du gaz est décrit par les équations de Navier-Stokes compressibles avec des conditions de glissement ;
- $\text{Kn} < 0.01$  : **régime hydrodynamique**. Les équations de Navier-Stokes avec les conditions d'adhérence à la paroi sont adaptées.

Les modèles simplifiés remplaçant l'équation de Boltzmann permettent généralement de retrouver ces limites fluides, à quelques différences de coefficients près. Le modèle BGK préserve ces limites, mais les différences dans les coefficients de transport entraînent toutefois une faiblesse majeure. L'application du développement de Chapman-Enskog à ce modèle donne pour la viscosité dynamique  $\mu$  et la conductivité thermique  $\kappa$  obtenus dans les relations (8) une expression linéaire en fonction du temps de relaxation :

$$\mu = p\tau \quad \text{et} \quad \kappa = \frac{5}{2} R p \tau ;$$

le nombre de Prandtl est alors

$$\text{Pr} = \frac{5}{2} R \frac{\mu}{\kappa} = 1.$$

Le modèle BGK donne ainsi un nombre de Prandtl fixé, différent des valeurs expérimentales ou obtenues avec certains potentiels de collision dans l'équation de Boltzmann : le modèle de sphères dures donne ainsi  $\text{Pr} = \frac{2}{3}$  pour les gaz monoatomiques ; les expériences donnent  $\text{Pr} = 0.71$  pour

l'air. Le modèle sous-estime donc la vitesse des transferts thermiques par rapport à celle des phénomènes hydrodynamiques, et surestime en conséquence les flux de chaleur. D'autres modèles plus élaborés permettent de corriger ce défaut, comme le modèle ES-BGK ("Elliptic-Statistical BGK") de Holway ([38] et [4] pour les principales propriétés mathématiques du modèle) ou le S-modèle de Shakhov [62] mais, outre le fait qu'ils peuvent souffrir d'autres défauts théoriques (comme une perte de la positivité de la fonction de distribution), ils ne sont pas toujours aisément exploitables numériquement. C'est pourquoi le modèle BGK, simple mais donnant des résultats relativement satisfaisants, reste très répandu pour la simulation numérique dans le cadre de la théorie cinétique (par ailleurs, des travaux assez récents ont montré qu'il était possible d'obtenir des nombres de Prandtl corrects avec le modèle BGK en utilisant des temps de relaxation dépendant de la vitesse microscopique  $\mathbf{v}$ , voir par exemple ce que propose Struchtrup dans [65]).

Cette résolution numérique des équations cinétiques fait appel à plusieurs catégories de méthodes. L'approche utilisée dépend essentiellement du régime de l'écoulement :

- en régime moléculaire libre, l'équation de Boltzmann est suffisamment simple (elle se réduit à une équation de transport) pour pouvoir établir des formules analytiques ;
- en régime raréfié, on utilise généralement une approche statistique, essentiellement avec la célèbre méthode Direct Simulation Monte-Carlo (DSMC) introduite par Graeme Bird en 1963 [14] ou une de ses variantes comme la méthode de Nanbu [56] ;
- en régime transitionnel, les méthodes statistiques deviennent coûteuses à mettre en oeuvre, en particulier pour les écoulements stationnaires, à cause de la densification du gaz qui oblige à utiliser un très grand nombre de particules numériques ; de plus, le bruit stochastique qui entache les résultats devient lui aussi très coûteux à réduire. On utilise alors plutôt des méthodes déterministes, comme des schémas Volumes Finis ou Galerkin Discontinu. La méthode "Boltzmann sur réseau" (Lattice-Boltzmann), qui repose sur un principe un peu différent et connaît actuellement un développement intense, appartient également à cette catégorie de méthodes ; cependant, cette méthode a été développée pour traiter des problèmes issus de la mécanique des fluides incompressible par une approche cinétique, et son extension récente à la dynamique des gaz en régime raréfié n'est valide que pour de très faibles nombres de Mach (généralement pour des problèmes de microfluidique) et ne peut donc être utilisée pour des problèmes de rentrée atmosphérique.

Dans le cadre de ses études sur la rentrée atmosphérique, le CEA-CESTA s'intéresse aux écoulements hypersoniques en régime transitionnel. Il dispose pour cela depuis une douzaine d'années d'un code d'aérodynamique raréfiée déterministe utilisant la méthode des Volumes Finis, mis au point par Luc Mieussens au cours de sa thèse [53]. Ce code permet de calculer une solution stationnaire du modèle BGK à vitesses discrètes pour une seule espèce gazeuse mono- ou polyatomique, sans dissociation des molécules. La géométrie peut être 2D ou 3D, plane ou axisymétrique.

Les modèles à vitesses discrètes consistent à considérer que les particules ne peuvent se déplacer que selon un nombre fini de vitesses microscopiques. C'est une première approximation du modèle cinétique continu, étudiée notamment par Renée Gatignol [36]. On peut également les voir comme une discrétisation partielle du modèle cinétique. Le modèle à vitesses discrètes que nous utilisons est le suivant : on définit un ensemble de multi-indices borné de  $\mathbb{N}^3$

$$\mathcal{K} = \left\{ k = (k_1, k_2, k_3) \in \mathbb{N}^3 \mid \forall 1 \leq i \leq 3, k_i \leq K_i \right\},$$

où  $K_1, K_2$  et  $K_3$  sont des bornes fixées. L'ensemble des vitesses discrètes  $\mathcal{V}$  est défini par

$$\mathcal{V} = \{\mathbf{v}_k, k \in \mathcal{K}\} \subset \mathbb{R}^3.$$

On remplace alors la distribution continue  $f(t, \mathbf{x}, \mathbf{v})$  par une distribution discrète  $f_{\mathcal{K}}(t, \mathbf{x}) = (f_k(t, \mathbf{x}))_{k \in \mathcal{K}}$ , où  $f_k(t, \mathbf{x})$  est une approximation de  $f(t, \mathbf{x}, \mathbf{v}_k)$ . Les intégrales continues sont remplacées par les intégrales discrètes

$$\langle g \rangle_{\mathcal{K}} = \sum_{k \in \mathcal{K}} g_k \omega_k,$$

où  $\omega_k$  est le poids d'intégration de la vitesse  $\mathbf{v}_k$ . Les moments et l'entropie discrets s'écrivent

$$\rho_{\mathcal{K}} = \langle \mathbf{m} f_{\mathcal{K}} \rangle_{\mathcal{K}} \quad H_{\mathcal{K}} = \langle f_{\mathcal{K}} \log f_{\mathcal{K}} \rangle_{\mathcal{K}}.$$

On résout alors le système

$$\forall k \in \mathcal{K}, \quad \frac{\partial f_k}{\partial t} + \mathbf{v}_k \cdot \nabla_{\mathbf{x}} f_k = \frac{1}{\tau} (\mathcal{E}_k - f_k),$$

où  $\mathcal{E}_k$  est un équilibre discret. La particularité essentielle du modèle de Luc Mieussens est d'utiliser comme équilibre une maxwellienne discrète permettant de respecter les lois de conservation (5), d'entropie (6) et la positivité au niveau discret, soit

$$\begin{cases} \frac{\partial}{\partial t} \langle \mathbf{m} f_{\mathcal{K}} \rangle_{\mathcal{K}} + \nabla \cdot \langle \mathbf{m} \otimes \mathbf{v} f_{\mathcal{K}} \rangle_{\mathcal{K}} = \mathbf{0}, \\ \frac{\partial}{\partial t} H_{\mathcal{K}} + \nabla \cdot \langle \mathbf{v} f_{\mathcal{K}} \log f_{\mathcal{K}} \rangle_{\mathcal{K}} \leq 0, \\ f_{\mathcal{K}}(t, \mathbf{x}) \geq 0 \quad \forall (t, \mathbf{x}). \end{cases}$$

L'idée de construction de l'équilibre discret  $\mathcal{E}_k$  est la suivante : en écrivant la maxwellienne continue sous la forme

$$M(f) = \exp(\boldsymbol{\alpha} \cdot \mathbf{m}),$$

avec  $\boldsymbol{\alpha} \in \mathbb{R}^5$  la variable entropique définie par

$$\boldsymbol{\alpha} = \left( \ln \left( \frac{\rho}{(2\pi RT)^{3/2}} \right) - \frac{|\mathbf{u}|^2}{2RT}, \frac{\mathbf{u}}{RT}, -\frac{1}{RT} \right)^T,$$

la relation (2.17) donne

$$\langle \mathbf{m} \exp(\boldsymbol{\alpha} \cdot \mathbf{m}) \rangle = \boldsymbol{\rho}.$$

La maxwellienne discrète est construite en utilisant cette relation au niveau discret : on la définit par

$$\forall k \in \mathcal{K}, \quad \mathcal{E}_k = \exp(\boldsymbol{\alpha} \cdot \mathbf{m}_k),$$

avec  $\mathbf{m}_k = \left( 1, \mathbf{v}_k, \frac{1}{2} |\mathbf{v}_k|^2 \right)^T$ , et où  $\boldsymbol{\alpha}$  est l'unique solution du système non linéaire

$$\langle \mathbf{m} \exp(\boldsymbol{\alpha} \cdot \mathbf{m}) \rangle_{\mathcal{K}} = \boldsymbol{\rho}_{\mathcal{K}}.$$

Le modèle obtenu est utilisé avec un schéma implicite linéarisé. Les techniques algorithmiques et les détails de la discrétisation sont abordées dans l'annexe du chapitre 1 ; ils sont également décrits précisément pour un modèle BGK réduit au chapitre 4.



Le code du CESTA utilise des maillages structurés par blocs, technique adaptée pour effectuer une parallélisation par décomposition de domaine en espace (MPI) et par boucles (OpenMP), ce qui permet des calculs performants sur les calculateurs dont dispose le CEA (supercalculateur TERA-100 depuis 2010). Cependant, ces calculs restent très coûteux à cause de la quantité de données utilisées par le code : la distribution  $f$  est en effet décrite dans un espace à 5 ou 6 dimensions (selon que l'espace physique est 2D ou 3D). En conséquence, les calculs stationnaires 3D sont encore limités, et les calculs instationnaires semblent pour le moment hors de portée. Les travaux exposés ici ont pour objectif de diminuer le coût des calculs. Deux évolutions principales ont été envisagées.

La première amélioration considérée a été un travail sur l'espace des vitesses discrètes  $\mathcal{V}$ . Le code du CEA utilise, comme la plupart des codes stationnaires, une grille de vitesses globale, identique en chaque point de l'espace physique, et conservée tout au long du calcul. En règle générale, on utilise une grille structurée cartésienne de pas constant. Or le nombre de points de vitesse qu'il est nécessaire d'utiliser dans cette grille dépend du nombre de Mach ; lors d'une rentrée atmosphérique, celui-ci peut atteindre des valeurs de plusieurs dizaines à cause de la vitesse élevée de l'engin de rentrée et de la basse température des couches d'air traversées ; la taille de la grille de vitesses devient alors considérable. Cela entraîne un coût de calcul prohibitif, tant au niveau CPU que de la place mémoire utilisée pour stocker l'ensemble des données, même pour un code fortement parallélisé utilisé sur de grands calculateurs.

L'une des pistes que nous avons donc étudiées afin de diminuer le coût des calculs est une méthode destinée à construire automatiquement une grille de vitesses adaptée à la configuration considérée, de façon à atteindre une précision satisfaisante des résultats tout en utilisant un nombre de points de vitesses raisonnable. L'idée principale consiste à estimer qu'une discrétisation fine de l'espace des vitesses n'est requise que dans certaines zones de cet espace, là où les distributions présentent un profil étroit en vitesse, et qu'il est possible d'utiliser des maillages de pas plus grand dans d'autres zones de l'espace sans altérer la précision des résultats. Il a notamment déjà été constaté, entre autres, qu'un resserrement du maillage autour de la vitesse nulle améliorerait la résolution du calcul. Ces constatations ont donné lieu à quelques tentatives de raffinements locaux de la grille des vitesses, notamment par Aristov dans les années 1970 pour des cas particuliers (voir [6]), mais jusqu'à présent il n'existait aucune procédure systématique de type Adaptative Mesh Refinement (AMR) pour construire automatiquement des grilles de vitesses localement raffinées sur une configuration de rentrée atmosphérique. Nous proposons ici un tel algorithme, qui a permis des gains substantiels en termes de place mémoire et de temps CPU dans nos simulations. Cette étude a fait l'objet d'un article publié au Journal of Computational Physics en 2013 [10], reproduit au chapitre 1.

L'autre voie que nous avons explorée est une amélioration du schéma en espace. Comme on l'a évoqué plus haut, la raréfaction du gaz autour du corps de rentrée crée à la surface de ce corps une couche limite cinétique, la couche de Knudsen. Cette couche limite très fine, bien plus que la couche limite visqueuse (son épaisseur est de l'ordre de quelques libres parcours moyens des molécules), a une influence déterminante sur le flux de chaleur à la paroi, ainsi que sur les coefficients aérodynamiques, grandeurs capitales à calculer pour étudier une rentrée atmosphérique ; il est donc très important de prendre en compte numériquement l'effet de cette couche limite. La méthode des Volumes Finis impose alors d'utiliser des maillages très fins en zone de proche paroi, résolvant cette couche limite, d'où un coût de calcul élevé.

Des études sur le transport linéaire des neutrons, notamment celle de Larsen, Morel et Miller [48] et les travaux plus récents de Guermond et Kanschat [37], ont montré que les schémas

Galerkin Discontinu pouvaient prendre en compte les effets d'une couche limite cinétique sans qu'elle soit résolue par le maillage. Il est alors possible d'utiliser ces schémas sur des maillages de taille raisonnable sans perdre d'information sur les bords du domaine de calcul. Les schémas présentant cette propriété sont dits "Asymptotic-Preserving" (AP), car elle est étroitement liée au fait que ces schémas sont capables de capturer la limite asymptotique du modèle cinétique associée au phénomène physique qui est à l'origine de la couche limite (dans le cas de l'équation de transport linéaire, il s'agit de la limite de diffusion ; dans celui de l'équation BGK appliquée à la dynamique des gaz, la couche de Knudsen est d'origine hydrodynamique, on considère donc une limite fluide).

Le principe des schémas AP est illustré sur la figure (6) : partant d'un problème cinétique continu dépendant d'un paramètre  $\varepsilon$  supposé petit, que l'on note  $P_\varepsilon$ , on construit un schéma discrétisant ce problème sur un maillage de pas  $h$ , noté  $P_\varepsilon^h$ . On appelle  $P_0$  la limite asymptotique du problème  $P_\varepsilon$ , c'est-à-dire la limite obtenue quand le paramètre  $\varepsilon$  tend vers 0. Le schéma est alors "AP" s'il préserve cette limite asymptotique, c'est-à-dire si le schéma limite  $P_0^h$  obtenu lorsque  $\varepsilon$  tend vers 0 est consistant avec  $P_0$ .

$$\begin{array}{ccc}
 P_\varepsilon^h & \xrightarrow{h \rightarrow 0} & P_\varepsilon \\
 \downarrow \varepsilon \rightarrow 0 & & \downarrow \varepsilon \rightarrow 0 \\
 P_0^h & \xrightarrow{h \rightarrow 0} & P_0
 \end{array}$$

FIGURE 6: Schéma asymptotiquement préservant.

Nous nous sommes donc intéressés à l'utilisation de schémas Galerkin Discontinu pour le modèle BGK stationnaire à vitesses discrètes, afin de déterminer si ces schémas présentent pour ce modèle une propriété AP permettant d'utiliser pour nos calculs des maillages ne résolvant pas la couche de Knudsen. En faisant l'hypothèse que près des parois le gaz est proche de l'équilibre thermodynamique, nous avons étudié par une décomposition de Hilbert l'équation BGK linéarisée autour de cet équilibre, avec les conditions de réflexion diffuse, afin d'établir rigoureusement ses limites fluides (le jeu d'équations de Stokes) et les propriétés de la couche de Knudsen, notamment les conditions aux limites de glissement obtenues à l'approche du régime transitionnel (chapitre 2). En appliquant un schéma Galerkin Discontinu à ce modèle linéarisé, nous montrons ensuite que cette limite fluide est préservée au moins pour le premier ordre de la décomposition de Hilbert (chapitre 3). Cette étude constitue une première étape vers l'analyse de l'écoulement à nombre de Mach fini, sans linéarisation de l'équation BGK, applicable à la simulation numérique de la rentrée atmosphérique mais beaucoup plus complexe à étudier.

Parallèlement à l'étude des schémas Galerkin Discontinu, nous avons travaillé sur une amélioration possible des conditions aux limites utilisées dans les schémas Volumes Finis. En effet, nous avons constaté que dans beaucoup de schémas classiques d'ordre 2 ou d'ordre supérieur, les conditions aux limites numériques ne sont pas adaptées et font dégénérer les résultats à l'ordre 1 sur les bords du domaine de calcul. Nous avons donc étudié les conditions à imposer pour

conserver un ordre élevé sur la totalité du domaine de calcul, dans les cas où cela est possible. Des tests numériques ont été effectués, afin d'analyser les performances de ces schémas et de les comparer à un schéma Galerkin Discontinu, sur un cas-test 1D (écoulement de Couette) pour lequel on a construit un modèle BGK réduit permettant d'utiliser un espace de vitesses discrètes 1D. Les résultats de cette étude sont présentés au chapitre 4.

# Chapitre 1

## Raffinement local des grilles de vitesses discrètes

*C'est un tort égal de pécher par excès ou par défaut.*

*Confucius - Livre des sentences*

**Abstract** : Most of deterministic solvers for rarefied gas dynamics use discrete velocity (or discrete ordinate) approximations of the distribution function on a Cartesian grid. This grid must be sufficiently large and fine to describe the distribution functions at every space position in the computational domain. For 3-dimensional hypersonic flows, like in re-entry problems, this induces much too dense velocity grids that cannot be practically used, for memory storage requirements. In this article, we present an approach to generate automatically a locally refined velocity grid adapted to a given simulation. This grid contains much less points than a standard Cartesian grid and allows us to make realistic 3-dimensional simulations at a reduced cost, with a comparable accuracy.

**Keywords** : rarefied flow simulation, hypersonic flows, re-entry problem, transition regime, BGK model, discrete velocity model

### 1.1 Introduction

The description of a flow surrounding a flying object at hypersonic speed in the rarefied atmosphere (more than 60 km altitude) is still a challenge in the atmospheric Re-Entry community [2]. When this flow is in a rarefied state, that is to say when the Knudsen number (which is the ratio  $\text{Kn} = \frac{\lambda}{L}$  between the mean free path  $\lambda$  of particle and a characteristic macroscopic length  $L$ ) is larger than 0.01, the flow cannot be accurately described by the compressible Navier-Stokes equations of Gas Dynamics. In this case, the kinetic theory has to be used. The evolution of the molecules of the gas is then described by a mass density distribution in phase space, which is a solution of the Boltzmann equation. In the transitional regime, this equation can be replaced by the simpler Bhatnagar-Gross-Krook (BGK) model.

In order to be able to compute parietal heat flux and aerodynamic coefficients in the range of 60-120 km, a kinetic description of the stationary flow is necessary.

The most popular numerical method to simulate rarefied flows is the Direct Simulation Monte Carlo method (DSMC) [14]. However, it is well known that this method is very expensive in transitional regimes, in particular for flows in the range of altitude we are interested in. The efficiency of DSMC can be improved by using coupling strategies (see [20, 31]) or implicit schemes (see [32, 57]), but these methods are still not very well suited for stationary computations. In contrast, deterministic methods (based on a numerical discretization of the stationary kinetic model) can be more efficient in transitional regimes. Up to our knowledge, there are few deterministic simulation codes specifically designed for steady flows. One of the most advanced ones is the 3D code of Titarev [67] developed for unstructured meshes. Another 3D code has been developed by G. Brook [18]. Other codes exist, but they are rather designed for unsteady problems, see for instance [1, 46] or the recent UGKS scheme developed by K. Xu and his collaborators [40, 71], which is an Asymptotic Preserving scheme for unsteady flows.

In our team, we developed several years ago a code to make 2D plane and axisymmetric simulations of rarefied flows based on the BGK model (see a description in [5, 51, 52]). This code has recently been extended to 3D computations, for polyatomic gases. Due to the physical model (polyatomic gases), the space discretization (block structured mesh), and the parallelization (space domain decomposition with MPI and inner parallelization with openMP), this code is rather different from the other existing 3D codes recently presented in the literature for the same kind of problems (the 3D code of Titarev [67] for example), even if space domain decompositions have already been used for unsteady simulation (see [46]).

All the codes designed for steady flows have a common feature : they are based on a “discrete ordinate” like approach, and use a global velocity grid. This grid is generally a Cartesian grid with a constant step size. The number of points of this grid is roughly proportional to the Mach number of the flow in each direction, and hence can be prohibitively large for hypersonic flows, even with parallel computers. To compute realistic cases (3D configurations with Mach number larger than 20), the velocity space discretization has to be modified in order to decrease CPU time and memory storage requirements. It has already been noticed that a refinement of the grid around small velocities can improve the accuracy and reduce the cost of the computation (especially for large Knudsen numbers in flows close to solid boundaries, see [67]). However, up to our knowledge, there is no general strategy in the literature that helps us to reduce the number of discrete velocities of a velocity grid for any rarefied steady flow, even if some works on adaptive velocity grids have already been presented : the first attempt seems to be [7] for a 1D shock wave problem, and recently, more general adaptive grid techniques designed for unsteady flows have been presented in [19, 24, 45, 47].

The main contribution of this article is to propose an algorithm for an automatic construction of a locally refined velocity grid that allows a dramatic reduction of the number of discrete velocities, with the same accuracy as a standard Cartesian grid. This algorithm uses a compressible Navier-Stokes pre-simulation to obtain a rough estimation of the macroscopic fields of the flow. These fields are used to refine the grid wherever it is necessary by using some indicators of the width of the distribution functions in all the computational domain. This strategy allows us to simulate hypersonic flows that can hardly be simulated by standard methods, since we are indeed able to apply our method to our kinetic code to simulate a re-entry flow at Mach 25 and for temperature and pressure conditions of an altitude of 90 km. In this example, the CPU time and memory storage can be decreased up to a factor 24, as compared to a method with a standard Cartesian velocity grid. Note that preliminary results have already been presented in [8] and [9].

The outline of this article is the following. In section 1.2, we briefly present the kinetic description of a rarefied gas. In section 1.3, we discuss the problems induced by the use of a

global velocity grid, and our algorithm is presented. Our approach is illustrated in section 1.4 with several numerical tests. To simplify the reading of the paper, the presentation of our simulation code is made in the appendix A.

## 1.2 Boltzmann equation and Cartesian velocity grid

### 1.2.1 Kinetic description of rarefied gases

In kinetic theory, a monoatomic gas is described by the distribution function  $f(t, \mathbf{x}, \mathbf{v})$  defined such that  $f(t, \mathbf{x}, \mathbf{v})d\mathbf{x}d\mathbf{v}$  is the mass of molecules that at time  $t$  are located in an elementary space volume  $d\mathbf{x} = (x, y, z)$  and have a velocity in an elementary volume  $d\mathbf{v}$  centered in  $\mathbf{v} = (v_x, v_y, v_z)$ .

Consequently, the macroscopic quantities as mass density  $\rho$ , momentum  $\rho\mathbf{u}$  and total energy  $E$  are defined as the five first moments of  $f$  with respect to the velocity variable, namely :

$$(\rho(t, \mathbf{x}), \rho\mathbf{u}(t, \mathbf{x}), E(t, \mathbf{x})) = \int_{\mathbb{R}^3} (1, \mathbf{v}, \frac{1}{2}|\mathbf{v}|^2) f(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}. \quad (1.1)$$

The temperature  $T$  of the gas is defined by the relation  $E = \frac{1}{2}\rho|\mathbf{u}|^2 + \frac{3}{2}\rho RT$ , where  $R$  is the gas constant defined as the ratio between the Boltzmann constant and the molecular mass of the gas.

When the gas is in a thermodynamical equilibrium state, it is well known that the distribution function  $f$  is a Gaussian function  $M[\rho, \mathbf{u}, T]$  of  $\mathbf{v}$ , called Maxwellian distribution, that depends only on the macroscopic quantities :

$$M[\rho, \mathbf{u}, T] = \frac{\rho}{(2\pi RT)^{\frac{3}{2}}} \exp(-\frac{|\mathbf{v} - \mathbf{u}|^2}{2RT}). \quad (1.2)$$

It can easily be checked that  $M$  satisfies relations (1.1).

If the gas is not in a thermodynamical equilibrium state, its evolution is described by the following kinetic equation

$$\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = Q(f), \quad (1.3)$$

which means that the total variation of  $f$  (described by the left-hand side) is due to the collisions between molecules (described by the right-hand side). The most realistic collision model is the Boltzmann operator but it is still very computationally expensive to use. In this paper, we use the simpler BGK model [13, 70]

$$Q(f) = \frac{1}{\tau} (M[\rho, \mathbf{u}, T] - f) \quad (1.4)$$

which models the effect of the collisions as a relaxation of  $f$  towards the local equilibrium corresponding to the macroscopic quantities defined by (1.1). The relaxation time is defined as  $\tau = \frac{\mu}{\rho RT}$ , where  $\mu$  is the viscosity of the gas. This definition allows to match the correct viscosity in the Navier-Stokes equations obtained by the Chapman-Enskog expansion. This viscosity  $\mu$  is usually supposed to fit the law  $\mu = \mu_{ref} \left( \frac{T}{T_{ref}} \right)^\omega$ , where  $\mu_{ref}$  and  $T_{ref}$  are reference viscosity

and temperature determined experimentally for each gas, as well as the exponent  $\omega$  (see a table in [14]).

The interactions of the gas with solid boundaries are described with the diffuse reflection model. Let us suppose that the boundary has a velocity  $\mathbf{u}_w = 0$  and temperature  $T_w$ . In the diffuse reflection model, a molecule that collides with this boundary is supposed to be re-emitted with a temperature equal to  $T_w$ , and with a random velocity normally distributed around 0. This reads

$$f(t, \mathbf{x}, \mathbf{v}) = M[\sigma_w, 0, T_w](\mathbf{v}) \quad (1.5)$$

if  $\mathbf{v} \cdot \mathbf{n}(\mathbf{x}) > 0$ , where  $\mathbf{n}(\mathbf{x})$  is the normal to the wall at point  $\mathbf{x}$  directed into the gas. The parameter  $\sigma_w$  is defined so that there is no normal mass flux across the boundary (all the molecules are re-emitted). Namely, that is,

$$\sigma_w = - \left( \int_{\mathbf{v} \cdot \mathbf{n}(\mathbf{x}) < 0} f(t, \mathbf{x}, \mathbf{v}) \mathbf{v} \cdot \mathbf{n}(\mathbf{x}) d\mathbf{v} \right) / \left( \int_{\mathbf{v} \cdot \mathbf{n}(\mathbf{x}) > 0} M[1, 0, T_w](\mathbf{v}) \mathbf{v} \cdot \mathbf{n}(\mathbf{x}) d\mathbf{v} \right). \quad (1.6)$$

There are other reflection models, like the Maxwell model with partial accommodation, but they are not used in this work.

### 1.2.2 Velocity discretization on Cartesian grid

In most of existing computational codes for the Boltzmann equation that use a deterministic method, the velocity variable is discretized with a *global* Cartesian grid, that is to say the same grid for every point of the space mesh. Consequently, it is necessary to compute a priori a velocity grid in which every distribution function that may appear in the computation can be resolved (see figure 1.1). This means that the grid must be :

- large enough to contain the main part of the distribution functions *for every position in the computational domain*
- fine enough to capture the core of the distribution functions *for every position in the computational domain*

The corresponding parameters (bounds and step of the grid) can be determined with the following idea : at each point  $\mathbf{x}$  of the computational domain, the macroscopic velocity  $\mathbf{u}(\mathbf{x})$  and temperature  $T(\mathbf{x})$  give information on the support of the distribution function  $f(\mathbf{x}, \cdot)$ . Indeed, if  $f(\mathbf{x}, \cdot)$  is “not too far” from its corresponding Maxwellian, as supposed in the BGK model, its support is centered around  $\mathbf{u}(\mathbf{x})$ , and is essentially localized between bounds that depend on  $\mathbf{u}(\mathbf{x})$  and  $T(\mathbf{x})$ , corresponding to the standard deviation of the Maxwellian ; that is to say  $f(\mathbf{x}, \cdot)$  is very small outside (see figure 1.2 where these bounds are  $\mathbf{u}(\mathbf{x}) - c\sqrt{RT(\mathbf{x})}$  and  $\mathbf{u}(\mathbf{x}) + c\sqrt{RT(\mathbf{x})}$  in 1D). When several distribution functions have to be discretized on the same grid, their supports are reasonably well approximated if the bounds are

$$v_{max}^\alpha = \max_{\mathbf{x} \in \Omega} \{u^\alpha(\mathbf{x}) + c\sqrt{RT(\mathbf{x})}\} \quad v_{min}^\alpha = \min_{\mathbf{x} \in \Omega} \{u^\alpha(\mathbf{x}) - c\sqrt{RT(\mathbf{x})}\}, \quad (1.7)$$

and if the grid step is

$$\Delta v = a \min_{\mathbf{x} \in \Omega} \sqrt{RT(\mathbf{x})}, \quad (1.8)$$

where  $\Omega$  is the computational domain (in space), and  $\alpha = x, y, z$  (see 1.3). In (1.7)-(1.8), the parameters  $c$  and  $a$  can be chosen as follows. For  $c$ , statistical argument suggest that values between 3 and 4 are needed, and so  $c = 3$  seems to be a good compromise between accuracy and computational cost. The parameter  $a$  allows us to adjust the grid step : it must be at most equal to 1 (which ensures that there are at least six points inside the core of every distribution function), but a smaller value might be necessary for an accurate computation.

To compute these bounds, it is necessary to a priori estimate the values of  $\mathbf{u}$  and  $T$  in the computational domain. A simple way is to use a compressible Navier-Stokes (CNS) pre-simulation (for which the CPU time is negligible as compared to the Boltzmann simulation). This simulation gives the fields  $\mathbf{u}^{CNS}$  and  $T^{CNS}$  in  $\Omega$ , and the bounds and the grid step can be defined by formulas (1.7)-(1.8) applied to these fields, that is to say

$$v_{max}^\alpha = \max_{\mathbf{x} \in \Omega} \{u^{\alpha, CNS}(\mathbf{x}) + c\sqrt{RT^{CNS}(\mathbf{x})}\} \quad v_{min}^\alpha = \min_{\mathbf{x} \in \Omega} \{u^{\alpha, CNS}(\mathbf{x}) - c\sqrt{RT^{CNS}(\mathbf{x})}\}, \quad (1.9)$$

and

$$\Delta v = a \min_{\mathbf{x} \in \Omega} \sqrt{RT^{CNS}(\mathbf{x})}. \quad (1.10)$$

However, the values of the bounds are mainly determined by the temperatures in the shock zone, which can reach thousands of Kelvin. One must be careful that, at high altitude and high speeds, numerical computations of Navier-Stokes equations can underestimate those temperatures, which would lead to inappropriate bounds.

It is quite clear that the size of the velocity grid increases with the Mach number. Indeed, a large Mach number implies large upstream velocities and large temperature in the shock wave, which lead to very large bounds (see (1.7)), while the temperature of the body remains small, thus leading to small grid step (see (1.8)). For realistic re-entry problems, this can lead to prohibitively large grids. For instance, for the flow around a cylinder of radius 0.1 m at Mach 20 and altitude 90 km, formulas (1.9) and (1.10) used with a Compressible Navier-Stokes pre-simulation lead to a velocity grid of  $52 \times 41 \times 41$  points and around 350 GB memory requirements with a coarse 3D mesh in space.

However, this is mainly due to the use of a Cartesian grid with a constant step. In order to decrease the size of the grid, it is attractive to refine it only wherever it is necessary, and to coarsen it elsewhere. In the following section, we show that this can be done automatically with our new algorithm.

This approach may not be well designed for cases where  $f$  is very far from its corresponding Maxwellian (like in shock interactions problem). However, for reentry problems, our assumption is realistic for transitional regime we are interested in.

### 1.3 An algorithm to define locally refined discrete velocity grids

Since we have to represent many distribution functions on the same grid, it is natural to refine the grid in the cores of these distributions, and to coarsen it in the tails (see figure 1.4). To achieve this goal, we define the concept of “support function”, and we use it to design an AMR (Adaptive Mesh Refinement) velocity grid.



### 1.3.1 The support function

At each point  $\mathbf{x}$  of the computational space domain  $\Omega$ ,  $\mathbf{u}(\mathbf{x})$  and  $T(\mathbf{x})$  give information on the “main” support of the distribution function  $f(\mathbf{x}, \cdot)$  in the velocity space : this support is centered at  $\mathbf{u}(\mathbf{x})$  and of size  $2c\sqrt{RT(\mathbf{x})}$  (see (1.7)). We define the *support function*  $\phi$  in the velocity space as follows : for every  $\mathbf{v}$  in the velocity space, we set  $\phi(\mathbf{v}) = \sqrt{RT(\mathbf{x})}$  if there exists in  $\Omega$  a  $\mathbf{x}$  such that  $\|\mathbf{v} - \mathbf{u}(\mathbf{x})\| \leq c\sqrt{RT(\mathbf{x})}$ , that is to say that  $\mathbf{v}$  is inside the support of a distribution, considered as a sphere of center  $\mathbf{u}(\mathbf{x})$  and of radius  $c\sqrt{RT(\mathbf{x})}$ . This function gives an incomplete mapping of the velocity domain (see figure 1.5). Now we want to extend this function to the whole velocity space, so that it can be used as a refinement criterion to design an optimal velocity grid.

Indeed, so far, this function is not defined for every  $\mathbf{v}$ , especially for large ones, as we cannot find for every  $\mathbf{v}$  an appropriate pair  $(\mathbf{u}(\mathbf{x}), T(\mathbf{x}))$  to match the definition of the support function. Moreover, this function might be multivalued, as there might be two different space points  $\mathbf{x}_1$  and  $\mathbf{x}_2$  with same velocity  $\mathbf{u}(\mathbf{x}_1) = \mathbf{u}(\mathbf{x}_2)$ , but with different temperatures  $T(\mathbf{x}_1) \neq T(\mathbf{x}_2)$ . Since our goal is to obtain for every  $\mathbf{v}$  the minimum size of the support of every distribution centered around  $\mathbf{v}$ , these two problems can be avoided as follows :

(a) we set

$$\phi(\mathbf{v}) = \min(\sqrt{RT(\mathbf{x}_1)}, \sqrt{RT(\mathbf{x}_2)}) \text{ if } \|\mathbf{v} - \mathbf{u}(\mathbf{x}_1)\| \leq c\sqrt{RT(\mathbf{x}_1)}$$

and

$$\|\mathbf{v} - \mathbf{u}(\mathbf{x}_2)\| \leq c\sqrt{RT(\mathbf{x}_2)} \text{ with } (\mathbf{u}(\mathbf{x}_1), T(\mathbf{x}_1)) \neq (\mathbf{u}(\mathbf{x}_2), T(\mathbf{x}_2)).$$

Indeed, the size of the grid around  $\mathbf{v}$  is constrained by the distribution centered on  $\mathbf{v}$  that has the smallest support, hence  $\phi$  must have the corresponding smallest possible value.

(b) if there is no  $\mathbf{x}$  such that  $\|\mathbf{v} - \mathbf{u}(\mathbf{x})\| \leq c\sqrt{RT(\mathbf{x})}$ , then  $\mathbf{v}$  is in the tail of every distribution function, and there is no reason to refine the grid there, so we set  $\phi(\mathbf{v})$  to its largest possible value  $\phi(\mathbf{v}) = \max \sqrt{RT(\mathbf{x})}$ .

These ideas lead to the following algorithm for an automatic construction of  $\phi$ . Note that below, the computational domain in space is supposed to be discretized with a mesh composed of cells numbered with three indices  $(i, j, k)$ . This is not a restriction of our algorithm.

*Algorithm 1.3.1* (Construction of the support function).

1. CNS velocity and temperature are stored in arrays  $\mathbf{u}(i, j, k)$  and  $T(i, j, k)$  for  $i, j, k = 1 : i_{max}, j_{max}, k_{max}$ .
2. construction of a (fine) Cartesian velocity grid (based on the previous global criterion (1.9)-(1.10)) : points  $\mathbf{v}(q)$  with  $q = 0 : q_{max}$ .
3. computation of the field  $\sqrt{RT}$ , stored in an descending order in the 1D array  $\psi(I)$ , for  $I = 1 : i_{max} \times j_{max} \times k_{max}$ , so that  $\psi(I) = \sqrt{RT((i, j, k)(I))}$ .
4. initialization of the array  $\phi(0 : q_{max}) = \max(\psi)$  on the velocity grid (one value per discrete velocity)
5. *do*  $I = 1 : i_{max} \times j_{max} \times k_{max}$  (loop on the cells of the space mesh)

do  $q = 0 : q_{max}$  (loop on the nodes of the velocity grid)

➤ if  $\mathbf{v}(q)$  is in the sphere of center  $\mathbf{u}((i, j, k)(I))$  and radius  $c\psi(I)$ , then it is in the support of a distribution function, and we set  $\phi(q) := \psi(I)$

This algorithm ensures that the array  $\phi$  satisfies the following property.

*Property 1.3.1.* For every  $q$  between 0 and  $q_{max}$ , we have :

$$\phi(q) = \min \left( \min_{i,j,k} \left\{ \sqrt{RT(i, j, k)} \text{ such that } \|\mathbf{v}(q) - \mathbf{u}(i, j, k)\| \leq c\sqrt{RT(i, j, k)} \right\}, \max_{i,j,k} \sqrt{RT(i, j, k)} \right)$$

This means that if we take a velocity  $\mathbf{v}(q)$  of the fine grid, then among all the distribution functions whose support contains  $\mathbf{v}(q)$ , one of them has a smallest support, and  $2c\phi(q)$  is the size of its support. This support function hence tells us how the fine Cartesian grid should be refined, or coarsened, around  $\mathbf{v}(q)$ .

The generation of the corresponding adapted grid is described in the following section.

*Remark 1.1.* It may happen that with such a strategy, the support function does not take sufficiently into account the support of the wall Maxwellian (for diffuse reflection conditions on a solid obstacle) : this Maxwellian is centered around  $\mathbf{u} = 0$  and has a temperature  $T_{wall}$  which is often the smallest temperature of the computational domain. This can be the case when the macroscopic CNS flow is not resolved enough close to the wall, or if the CNS equations are solved with slip boundary conditions. In this case, it is safer to add the element  $\phi(q_{max} + 1) := \sqrt{RT_{wall}}$  in  $\phi$  and to add the corresponding velocity  $\mathbf{v}(q_{max} + 1) := 0$  in the fine grid.

### 1.3.2 AMR grid generation

The idea is to start with a unique cell defined by the bounds of the fine velocity grid (that is the full velocity domain), and then to apply a recursive algorithm : each cell is cut if its dimensions are larger than the minimum of  $a\phi$  in the cell. The algorithm is the following :

*Algorithm 1.3.2* (Recursive refinement of a cell  $\mathcal{C}$ ).

1. compute the minimum of the field  $a\phi$  in the cell  $\mathcal{C}$ , that is to say the minimum of the elements of  $a\phi$  that have the same indices as the discrete velocities of the fine grid included in this cell  $\mathcal{C}$  :

$$m := a \min \{ \phi(q), \text{ such that } \mathbf{v}(q) \in \text{cell } \mathcal{C} \}$$

2. if one edge of  $\mathcal{C}$  is larger than  $m$ , then the cell is cut into 8 new subcells on which the refinement algorithm is applied, recursively.
3. else, the cell is kept as it is, and the cell and its vertices are numbered.

At the end of the algorithm, the final grid satisfies the following property : every cell has a size smaller than  $a \times$  the minimum of the support function in the cell.

### 1.3.3 An example

We anticipate on the numerical results that will be discussed in section 1.4 to illustrate the previous algorithms.

The test case is a 2D steady flow around a cylinder at Mach 20 (see the geometry in figure 1.6 and the physical parameters in section 1.4.2). A CNS pre-simulation and the use of formula (1.9)-(1.10) with parameters  $c = 4$  and  $a = 2$  give us a fine velocity grid with  $52 \times 41$  points, see figure 1.7. Algorithms 1.3.1 and 1.3.2 applied to the CNS fields give the support function and the AMR velocity grid shown in figure 1.8. The AMR velocity grid has 529 points, which is one fourth as small as the original fine Cartesian grid.

Note that the AMR grid is refined around the zero velocity and the upstream velocity : these velocities correspond to the flow close to the solid boundary and to the upstream boundary, where the temperature is low, thus where the distribution functions are narrow, and hence where, indeed, we need a fine velocity resolution. At the contrary, the grid is coarse close to its boundaries : the discrete velocities are very large there and are all in the tails of all the distribution functions, where, indeed, we do not need a fine resolution.

The accuracy of the computation on this AMR grid is studied in section 1.4.

### 1.3.4 Discrete Velocity Approximation on the AMR velocity grid

When one wants to transform a standard discrete velocity method based on a Cartesian grid to a method using our AMR grid, two points have to be treated carefully : the computation of the moments of  $f$ , and the approximation of the collision operator. In this paper, we only use the BGK collision model of the Boltzmann equation : therefore its approximation (based on the conservative approach of [52]) reduces to the problem of a correct approximation of the moments of  $f$ , as it is described below.

First note that the standard discrete velocity approach consists in replacing the kinetic equation (1.3) by the following discrete kinetic equation

$$\partial_t f_q + \mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q = Q_q(f), \quad (1.11)$$

where  $f_q(t, x)$  is supposed to be an approximation of  $f(t, x, \mathbf{v}_q)$  for every discrete velocity  $\mathbf{v}_q$ , with  $q = 0$  to  $q_{max}$ . The discrete distribution function is  $f = (f_q(t, x))$ , and  $Q_q$  is the discrete collision operator. The moments of the discrete distribution function are obtained by using a quadrature formula to replace (1.1) by

$$(\rho(t, \mathbf{x}), \rho \mathbf{u}(t, \mathbf{x}), E(t, \mathbf{x})) = \sum_{q=0}^{q_{max}} (1, \mathbf{v}_q, \frac{1}{2} |\mathbf{v}_q|^2) f_q(t, \mathbf{x}) \omega_q,$$

where  $(\omega_q)$  are the weights of the quadrature.

For the BGK model (1.4), the conservative approximation of [52] gives the following discrete BGK collision operator

$$Q_q(f) = \frac{1}{\tau} (M_q[\rho, \mathbf{u}, T] - f_q),$$

where  $M_q[\rho, \mathbf{u}, T]$  is the discrete Maxwellian that has the same discrete moments as  $f$ . It can be shown that there exists  $\alpha \in \mathbb{R}^5$  such that  $M_q[\rho, \mathbf{u}, T] = \exp(\alpha \cdot m(\mathbf{v}_q))$ , that is to say,  $\alpha$

is the unique solution of the nonlinear system

$$\sum_{q=0}^{q_{max}} m(\mathbf{v}_q) \exp(\alpha \cdot m(\mathbf{v}_q)) \omega_q = \begin{pmatrix} \rho \\ \rho \mathbf{u} \\ E \end{pmatrix},$$

where we note  $m(\mathbf{v}) = (1, \mathbf{v}, \frac{1}{2}|\mathbf{v}|^2)^T$ . This system can be solved by a Newton algorithm (see details in [52] and a more economic version of the algorithm due to Titarev in [67]).

Consequently, to apply this approach to our AMR velocity grid, we just have to chose a quadrature formula. First, we propose the standard  $\mathbb{Q}_1$  bilinear interpolation of  $f$  on each cell of the AMR grid (by using the four corners of the cell). In this case, the quadrature points are the nodes of the grid, and so are the discrete velocities that are used in the transport term. The quadrature weights are

$$\omega_q = \frac{1}{2^d} \sum_{\text{cell } \mathcal{C}_l \ni \mathbf{v}_q} |\mathcal{C}_l|,$$

where  $d = 3$  in 3D or  $2$  in 2D, and  $|\mathcal{C}_l|$  is the volume (or the area in 2D) of  $\mathcal{C}_l$ .

The number of discrete velocities can be decreased if we take a simpler  $\mathbb{P}_0$  (constant per cell) interpolation. Here, the quadrature points are the centers of the cells, and so are the discrete velocities  $\mathbf{v}_q$  that are used in the transport term. The weights are simply the volume (or the area) of the cells :  $\omega_q = |\mathcal{C}_q|$ . Note that the number of discrete velocities that are used with this approach is equal to the number of cells of the AMR grid, which is smaller than the number of nodes, see section 1.4.2. We advocate the use of this latter approach, since it allows to save a lot of computer memory, and since we observed with our tests that the  $\mathbb{Q}_1$  interpolation does not give more accurate results.

*Remark 1.2.* Our AMR velocity grids are often very coarse at their boundaries (since velocities are very large there, and are in the tails of all the distribution functions). Consequently, passing from  $\mathbb{Q}_1$  to a  $\mathbb{P}_0$  grid may lead to a grid which is not large enough. Indeed, if the  $\mathbb{Q}_1$  grid is of length  $L$  (in the  $x$  direction, for instance) and its outer cells are of length  $l$ , then the length of the  $\mathbb{P}_0$  grid now is  $L - l$ . If  $l$  is large, the  $\mathbb{P}_0$  grid is not large enough. In that case, it is thus safer to modify step 2 of Algorithm 1.3.1 by increasing the bounds of the fine Cartesian grid :  $v_{max}^\alpha$  and  $v_{min}^\alpha$  are replaced by  $v_{max}^\alpha + \Delta v$  and  $v_{min}^\alpha - \Delta v$  for example, where  $\Delta v$  is the step of the fine grid.

### 1.3.5 AMR grid generation for axisymmetric flows

Many interesting 3D flows have symmetry axis (like flows around axisymmetric bodies with no incidence). In that case, if we assume that the symmetry axis is aligned with the  $x$  coordinate, it is interesting to write the kinetic equation in the cylindrical coordinate system  $(x, r, \varphi)$ , since the distribution function  $f$  is independent of  $\varphi$ , and we get

$$\partial_t f + v_x \partial_x f + \zeta \cos \omega \partial_r f - \frac{\zeta \sin \omega}{r} \partial_\omega f = Q(f), \quad (1.12)$$

where  $\zeta$  and  $\omega$  are the cylindrical coordinates of the velocity in the local frame  $(e_r, e_\varphi)$ , that is to say  $v \cdot e_r = \zeta \cos \omega$  and  $v \cdot e_\varphi = \zeta \sin \omega$ .

If the velocities of the upstream flow and of the solid boundaries have no component in the  $\varphi$  direction, then  $f$  is even with respect to  $\omega$ . Indeed, using the Galilean invariance of

equation (1.12), we apply the symmetry with respect to the  $(x, z)$  plane to the equation to get that  $f(t, x, r, v_x, \zeta, -\omega)$  is also a solution.

In this case, the discrete velocity grid must be constructed for variables  $(v_x, \zeta, \omega)$ , where  $\zeta$  is non negative,  $\omega$  is in  $[0, \pi]$ , and  $v_x$  can take any value. The construction of this grid is slightly different from the Cartesian case. We briefly comment on these differences in this section.

First, bounds in  $v_x$  and  $\zeta$  directions can be easily obtained and we set

$$\begin{aligned} v_{x,max} &= \max_{\Omega} \{u_x^{CNS} + c\sqrt{RT}\}, & v_{x,min} &= \min_{\Omega} \{u_x^{CNS} - c\sqrt{RT}\}, \\ \zeta_{max} &= \max_{\Omega} \{u_r^{CNS} + c\sqrt{RT}\}, & \zeta_{min} &= 0, \end{aligned} \quad (1.13)$$

where  $u_x$  and  $u_r$  are the macroscopic velocities in the axial and radial directions. Since  $\omega$  is a bounded variable, there is nothing else to do at this level.

The steps  $\Delta v_x$  and  $\Delta \zeta$  of the initial grid are obtained by (1.10) for  $v_x$  and  $\zeta$ . However, the choice of the step in the  $\omega$  direction is not obvious. Indeed, this variable has no link with the width of the distribution. Moreover, the grid in this direction must be fine enough to capture the derivative  $\partial_{\omega} f$ , while in the other directions, only some moments have to be captured. After several tests (and from the experience made in [52]), it seems that a grid step  $\Delta \omega$  such that there are 30 points uniformly distributed in  $[0, \pi]$  is a good compromise between cost and accuracy.

To refine/coarsen this initial cylindrical grid, we choose to use the procedure described in the steps of sections 1.3.1 and 1.3.2 in the plane  $(v_x, \zeta)$  to generate a two-dimensional AMR grid. Indeed, for the same reasons as mentioned above, there is no reason to refine or coarsen the grid in the  $\omega$  direction. Then, the complete grid is obtained by a rotation of the two-dimensional AMR grid in the  $\omega$  direction, with the same step  $\Delta \omega$  as in the initial grid. This procedure is simple and turned out to be very efficient.

Since the AMR grid is uniform in the  $\omega$  direction, the discretization of (1.12) is made with the approximations derived in [52].

## 1.4 Numerical results

### 1.4.1 The 3D kinetic code

Our 3D kinetic code is written in Fortran, and can deal with 2D planar flows, or 2D axisymmetric flows as well as 3D flows on multiblock curvilinear structured grids. A steady state solution of the BGK equation for monoatomic or polyatomic gases is computed with a linearized implicit finite volume scheme, based on a velocity discretization which is conservative and entropic (see [51]). In axisymmetric cases, a specific scheme (named T-UCE) is used in order to ensure a conservative and entropic discretization of the transport operator (see [52]).

In order to deal with high computational costs (in 3D configurations for example), an hybrid parallel implementation is used : a domain decomposition in the position space used with the library routines specified by the Message Passing Interface (MPI) can deal with a large number of mesh cells. Note that this is the opposite of the strategy used in [67] (which uses a domain decomposition in the velocity space). Moreover, the OpenMP library is used with 8 threads for loops computations that are local w.r.t the position (computation of the moments, Maxwellian,

etc.), thus large numbers of velocity points can be reached. We also use OpenMP for loops that are local w.r.t the velocity, like the computation of the transport terms, for instance. The communication between different domains are treated as explicit boundary conditions. This code has been run on the super-computer TERA-100 of the CEA (that has around 140 000 cores Intel Xeon 7500), by using at most 1600 cores.

Since a general description of our code is not the main subject of this paper, we do not give more details here. However, for completeness, note that interested readers can find a more detailed description of the code in the appendix.

The distribution is initialized by using a CNS pre-simulation (also used for the AMR strategy) : the number of iterations before reaching convergence is then considerably decreased (see section 1.4.5).

For the 2D plane examples shown below, we use a simpler version of the code, called CORBIS. This code is based on the same tools but makes use of the reduced distribution technique to reduce the velocity space to 2D only. This code only uses OpenMP directives to work on parallel computers (see [5]). It has been used on a SMP node of 48 cores AMD-Opteron-8439-SE.

For the compressible Navier-Stokes simulations, we use a 3D code developed by the CEA-CESTA for hypersonic flows at moderate to low altitudes. The same techniques as in our 3D kinetic code are used : a finite volume linearized implicit scheme with multiblock structured meshes. On solid walls, no-slip and isothermal boundary condition are used ( $u = 0$  and  $T = T_{wall}$ ).

### 1.4.2 A 2D plane example

Here, we illustrate our approach with the simulation of a steady flow over a infinite cylinder of radius 0.1m at Mach 20, for density and pressure of the air at an altitude of 90 km. The gas considered here is argon (molecular mass=  $6.663 \times 10^{-24}kg$ ). Namely, we have  $\rho = 3.17 \times 10^{-6}kg/m^3$ ,  $u = 5.81 \times 10^3m/s$ , and  $T = 242.4K$ . The temperature of the wall is 293K. The downstream flow is ignored, and outflow boundary conditions are used at the boundary of the right side of the domain. Since the flow is invariant along  $z$ , the simulation is made with the 2D plane model in the plane  $(x, y, 0)$ . Finally, since the flow is symmetric with respect to the line  $y = 0$ , the computation is made on the upper part on this plane only (see figure 1.6). See appendix A.2 for the implementation of the boundary conditions in the code.

The space mesh uses  $50 \times 50$  cells, with a refinement such that the first cell at the solid boundary is smaller than one fifth of the mean free path at the boundary.

A CNS pre-simulation and the use of formula (1.9-1.10) with parameters  $c = 4$  and  $a = 2$  give us a fine velocity grid with  $52 \times 41$  points (bounds  $\pm 11\,500$  and  $\pm 8\,900$  for  $v_x$  and  $v_y$  with a step of 450, in  $m.s^{-1}$ ), see figure 1.7. Algorithms 1.3.1 and 1.3.2 applied to the CNS fields give the support function and the AMR velocity grid shown in figure 1.8. The AMR velocity grid has 529 points, which is one fourth as small as the original fine Cartesian grid. This gain can be further increased if we define the discrete velocities as the centers of the cells of the AMR grid rather than its vertices. This gives 316 discrete velocities, hence with a gain of 6.7 times instead of 4.

First, we compare the CPU time required by the code CORBIS with the different velocity grids. While the number of iterations to reach steady state is approximately the same with both

grids, the CPU time required by the original fine Cartesian velocity grid is around 7 times as large as with the new AMR grid, which is almost the same ratio as the ratio of the number of discrete velocities. The memory required with the uniform grid method is around 170 MB whereas with the use of the AMR grid, only 25 MB of memory storage is required. This shows that the new AMR grid leads to a real gain both in memory storage and CPU time.

Then, we compare the accuracy of the results with the two grids for the macroscopic quantities. We compute the normal component of the heat flux to the boundary, which is a quantity of paramount importance in aerodynamic simulations : we find a maximum relative difference lower than 5%, which is reasonably small (see the profile of this flux on figure 1.9). We also compute the differences for the density, temperature and pressure in the whole computational domain :

- the mean quadratic relative difference over all the cells of the computational domain is 5% for the density, 1% for the temperature, 0.6% for the horizontal velocity, and 1.2% for the vertical velocity ;
- the maximum relative difference on each cell of the computational domain is 40% for the density, 69% for the temperature, and 157% for the velocity.

The maximum relative differences are observed at the solid wall for the density and the velocity, and in the upstream flow for the temperature.

This difference is quite large, and can be explained as follows. The smallest cells of our AMR velocity grid (that are around small velocities, like velocities at the solid wall) turn out to be smaller than the cells of the Cartesian grid (size 330 instead of 450). This means that our results with the AMR grid are probably *more accurate* than the results of the Cartesian grid. Consequently, the Cartesian grid results should not be considered as a reference for this comparison.

To confirm this analysis, we make a new simulation with a Cartesian grid with a uniform step of 330 (like the smallest step of the AMR grid). We observe that relative differences between the Cartesian grid and the AMR grid are much smaller :

- the mean quadratic relative difference over all the cells of the computational domain is 2% for the density, 0.5% for the temperature, 0.3% for the horizontal velocity, and 0.7% for the vertical velocity ;
- the maximum relative difference on each cell of the computational domain is 12% for the density, 13% for the temperature, 80% for the velocity.

The maximum relative difference is still too large for the velocity (80% at the boundary), but this quantity is very small in this zone and probably requires smaller velocity cells (that is a smaller parameter  $a$  for both grids). However, this inaccuracy does not deteriorates the results on the other quantities, in particular for the heat flux at the boundary. Indeed, the comparison of the heat fluxes is even better, since we find a relative difference lower than 2.5%, which is excellent. The comparison in terms of CPU time and memory storage is of course more favorable to the AMR grid here.

Note that we also did the same computations with an finer AMR grid obtained with the parameter  $a = 1$  : the difference with  $a = 2$  is less than 2% for the heat flux at the solid wall, which is quite good. The value  $a = 2$  is then clearly sufficient. However, if one is interested with the flow velocity at the boundary, we advocate to use  $a = 1$ , since the differences here are around 15%.

### 1.4.3 A 2D axisymmetric example

Here we consider the flow around a sphere of radius 0.1 m, at Mach 20, for density and pressure of the air at an altitude of 90 km (see figure 1.10). The gas is air (molecular mass =  $4.81 \times 10^{-26}$  kg), considered as a diatomic gas. Namely, the upstream density and pressures are  $\rho = 3.17 \times 10^{-6} \text{ kg/m}^3$ , and  $p = 0.16 \text{ N.m}^{-2}$ , and the temperature of the surface of the sphere is  $T = 280 \text{ K}$ . The downstream flow is ignored, and outflow boundary conditions are used at the boundary of the right side of the domain. Since the flow is rotationally symmetric, the simulation is made with the 2D axisymmetric model in the plane  $(x, y, 0)$ . Finally, since the flow is symmetric with respect to the line  $y = 0$ , the computation is made on the upper part on this plane only.

The space mesh uses  $70 \times 50$  cells, with a refinement such that the first cell at the solid boundary is smaller than one fifth of the mean free path at the boundary. A CNS pre-simulation and the use of formula (1.13) with parameter  $c = 4$  give us the following velocity bounds  $-8700$  and  $11000$  for  $v_x$ ,  $0$  and  $7500$  for  $\zeta$ , and the bounds for  $\omega$  are always  $0$  and  $\pi$ , as explained in section 1.3.5.

All the following simulations were made with 140 domains and 4 OpenMP threads per domain.

First, a fine uniform cylindrical velocity grid is obtained with (1.10) and  $a = 2$  : this gives  $65 \times 33 \times 31 = 66495$  discrete velocities. The steady state is reached after 1828 iterations, in  $5476 \text{ s}$ . The corresponding temperature and velocity fields are shown in figure 1.11.

Then a cylindrical AMR grid is generated (with  $a = 2$ ) : it gives 7691 discrete velocities (with the  $\mathbb{Q}_1$  points), the steady state is reached after 1899 iterations, in  $710 \text{ s}$ . The support function and AMR grid are shown in figure 1.12. Here, the gain factors in memory and in CPU time with respect to the uniform grid are around 9.

Now, we use the same AMR grid, but with the  $\mathbb{P}_0$  points (that is to say the centers of the cells), which gives 6900 discrete velocities only. The number of iterations is close (1796), and the CPU time is  $562 \text{ s}$ . Here, the gain factors in memory and CPU time, with respect to the uniform grid, are around 10.

Finally, we compare in figure 1.13 the normal heat flux at the boundary for the uniform grid and the AMR grid generated by using the CNS fields ( $\mathbb{Q}_1$  version). As in the 2D example, we find that the fluxes are very close (with a relative difference which is less than 1%). We get the same results with the  $\mathbb{P}_0$  version.

Consequently, the AMR grid ( $\mathbb{P}_0$  version) is the most efficient here : it gives the same results as the fine uniform grid, for computational and memory cost which is one ninth as small.

*Remark 1.3.* Before discussing a full 3D case, we briefly mention another experiment we have done here. We have tried to see if the use of a CNS pre-simulation could be avoided by using a direct estimation of the extreme values of the velocity and the temperature given by the Rankine-Hugoniot relations. Indeed, before the computation, we know the upstream values  $\mathbf{u}_{up}$  and  $T_{up}$ , and we can compute the downstream values  $(\mathbf{u}_{down}, T_{down}) = RH(\mathbf{u}_{up}, T_{up})$  by the Rankine-Hugoniot relations for a stationary shock. Then we have three set of different values : upstream, downstream, and wall values  $(\mathbf{u}_{wall} = 0, T_{wall})$ , and we can use these three sets in our algorithm 1.3.1 instead of the CNS fields  $(\mathbf{u}^{CNS}, T^{CNS})$ .



With the same test case, this strategy gives a grid with 3962 points only (with  $a = 2$ ), which is half as less as with the CNS fields. The support function and AMR grid are shown in figure 1.14 : of course, since we only have three values for the macroscopic fields here, the grid is not as smooth as with the CNS fields.

Moreover, the convergence to steady state is much slower (since we cannot use the CNS solution as an initial state, see section 1.4.5 for a comment on the initialization). More important, the accuracy observed on the heat flux is not as good as with the CNS fields : the difference with the fine uniform grid is here around 8% instead of 1%.

Consequently, even if strategy is simple and does not require a CNS solver, it seems to be not accurate enough.

#### 1.4.4 A 3D example

Here, we consider the flow around a 3D object composed of cone with a spherical nose, with no incidence (see figure 1.15). The physical parameters are the same as in section 1.4.3. The space mesh uses 50 000 cells, with 50 in the direction normal to the surface. Again, the downstream flow is ignored, and outflow boundary conditions are used at the boundary of the right side of the domain.

A CNS pre-simulation and the use of formula (1.13) with parameter  $c = 4$  and  $a = 2$  give us a uniform Cartesian grid with  $65 \times 33 \times 33 = 70\,785$  discrete velocities. This is much too large to make the simulation (the memory storage itself is huge). Then we only do the simulation with the AMR grid ( $\mathbb{P}_0$  version) generated by our algorithm which has only 2956 points.

The gain factor in memory storage is 22. The steady state is reached in 3424 iterations and 3566s with 400 domains and 4 threads per domain.

To estimate the gain in CPU that would be obtained if the uniform Cartesian grid could be used, we use the following remark : on the previous tests (2D plane and axisymmetric), the number of iterations to reach the steady state is almost the same for the uniform and the AMR velocity grids. Then we make a single iteration with the uniform Cartesian grid (same domain decomposition) and we multiply the CPU time obtained for this iteration by the number of iterations required with the AMR grid. We find a CPU time which is 24 times as large as with the AMR grid, which is quite important. This CPU gain is larger than the gain observed for the memory storage.

In figure 1.17, we show the uniform Cartesian grid and the AMR grid. In figure 1.16, we show the pressure field computed by the code with the AMR grid, and the heat flux along the surface is shown in figure 1.18.

#### 1.4.5 Initialization with the CNS results

Since we use a CNS pre-simulation to build our discrete velocity grid, it is interesting to use it also to initialize the BGK computation, what was done in the previous simulations. It helps to reach the steady state more rapidly, as it is shown in the following example.

We take the same test case as in section 1.4.3, with the uniform velocity grid. In a first computation, the distribution is initialized with the (Maxwellian) upstream flow (this is the

standard initialization), and the scheme converges to the steady state in 3 549 iterations. Then we do the same computation, but now the initial state is the Maxwellian distribution computed with the macroscopic variables that are given by the steady CNS solution. Then the steady state is reached in 1 822 iterations only.

A similar gain (half as many iterations as with the standard initialization) is obtained in all our test cases.

## 1.5 Conclusion and perspectives

In this paper, we have proposed a method to refine and coarsen a global velocity grid of a discrete velocity approximation of a kinetic equation. It is based on a criterion called “the support function” that links the local size of the velocity grid to the macroscopic temperature and velocity of the flow. Our algorithm uses the macroscopic fields given by a compressible Navier-Stokes pre-simulation to automatically generate an optimal velocity grid.

This approach has been tested in a 3D computational code (and 2D plane and axisymmetric versions) which uses the BGK model of rarefied gas dynamics. Typical test cases in hypersonic re-entry problems (for simplified geometries) have been used. We observed that using our algorithm allows important gains both in CPU time and memory storage, up to a factor 24. This allows to make simulations that are hardly possible with standard grids, even on super computers.

Note that this work might be extended to unsteady flow simulations : we could indeed modify the velocity grid at different time steps according to our refinement algorithm. This might also be used during the iterations of our iterative solver for a steady simulation. However, this would require to interpolate the solution between two successive grids, and this might lead to an important increase of the CPU time. This clearly requires more investigations to get an efficient method.

An obvious, and straightforward, extension of this work will be the use of modified BGK models (like ES or Shakhov models), in order to have a model with a correct Prandtl number.

Moreover, we will try to investigate non isotropic AMR velocity grids by taking into account translational kinetic temperatures in different directions, that can be obtained by the stress tensor computed by the CNS simulation. It could also be interesting to refine the grid so as to capture the discontinuity of the distribution function in the velocity space, but this will require a rather different approach.

Our main short term project is to improve the physical accuracy of our code by implementing a simple BGK-like model for multi-species reactive flows. An application of our method to the full Boltzmann collision operator is less obvious but might also be explored.

## Acknowledgment

Experiments presented in this paper were carried out using the PLAFRIM experimental testbed, being developed under the Inria PlaFRIM development action with support from LABRI and IMB and other entities : Conseil Régional d’Aquitaine, FeDER, Université de Bordeaux and CNRS (see <https://plafirm.bordeaux.inria.fr/>).

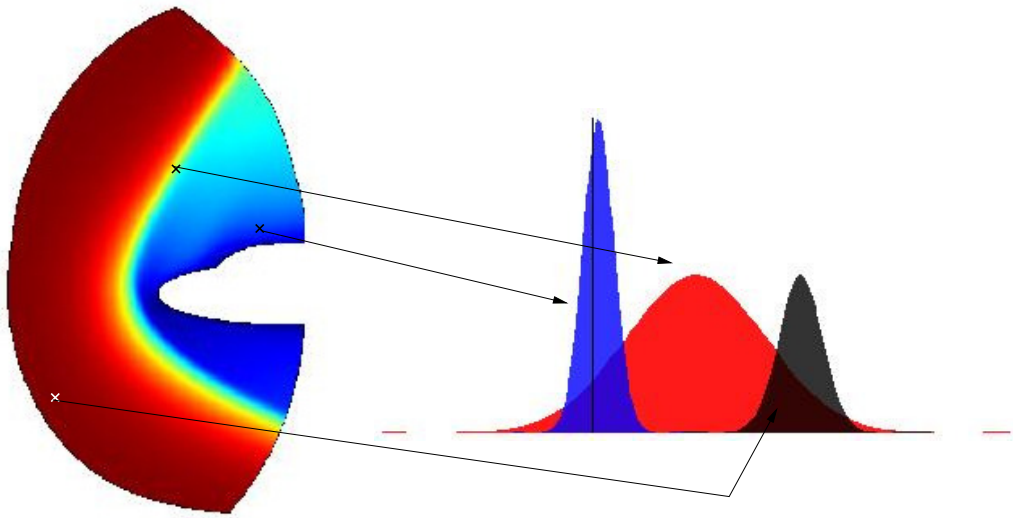


FIGURE 1.1: Three distribution functions in different space points of a computational domain for a re-entry problem

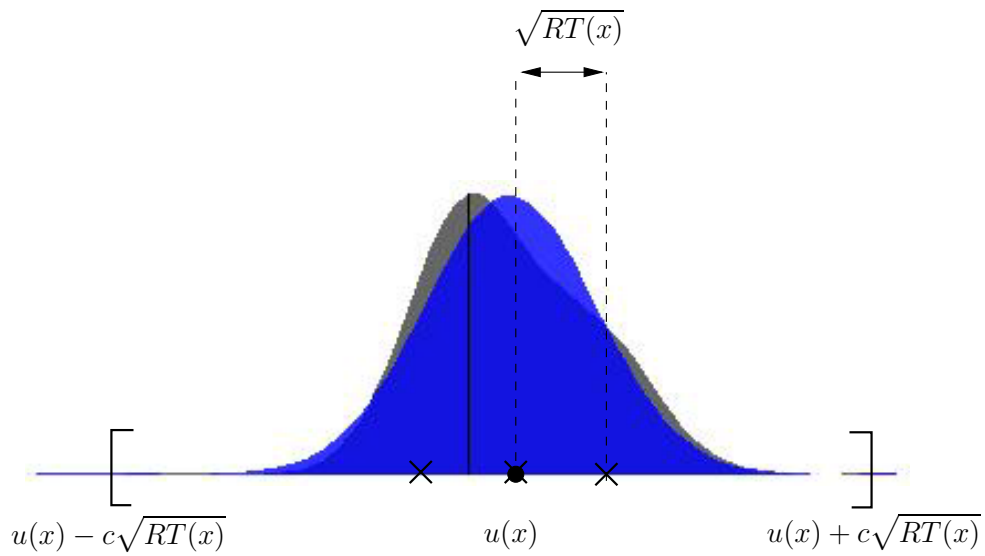


FIGURE 1.2: Support of a distribution function  $f(x, \cdot)$  : the distribution (in black), its corresponding Maxwellian (in blue), and the corresponding support based on the bulk velocity  $u$  and the temperature  $T$ .

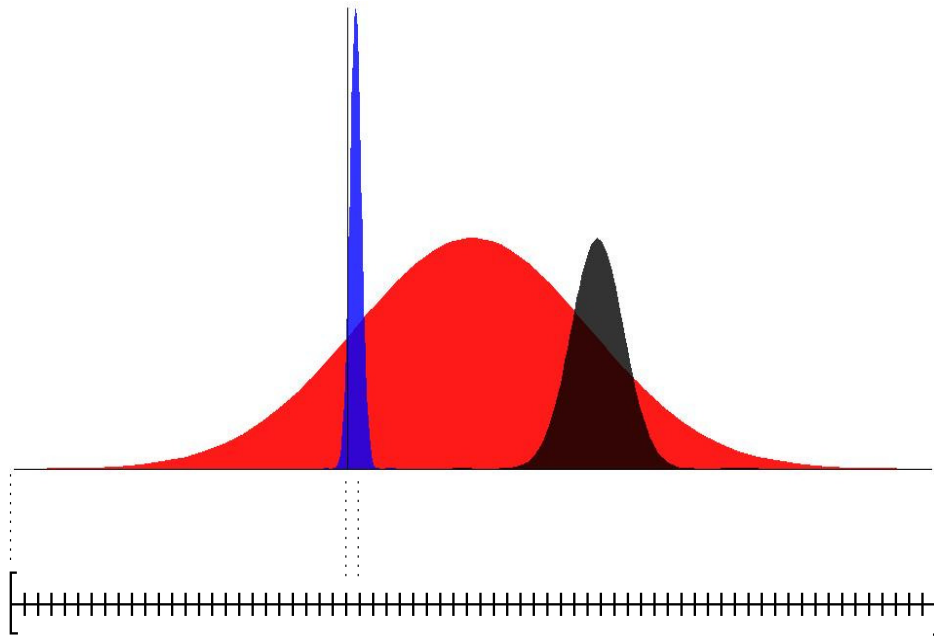


FIGURE 1.3: Three different distribution functions and a global grid that satisfies (1.7)-(1.8) : the bounds are given by the largest distribution, the step is given by the narrowest distribution.

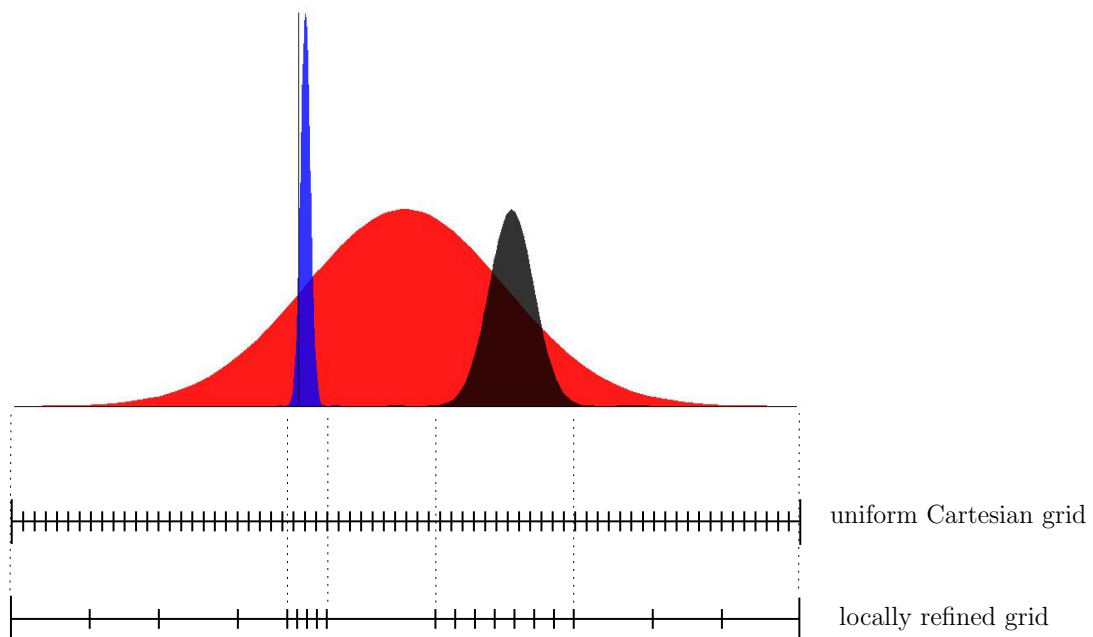
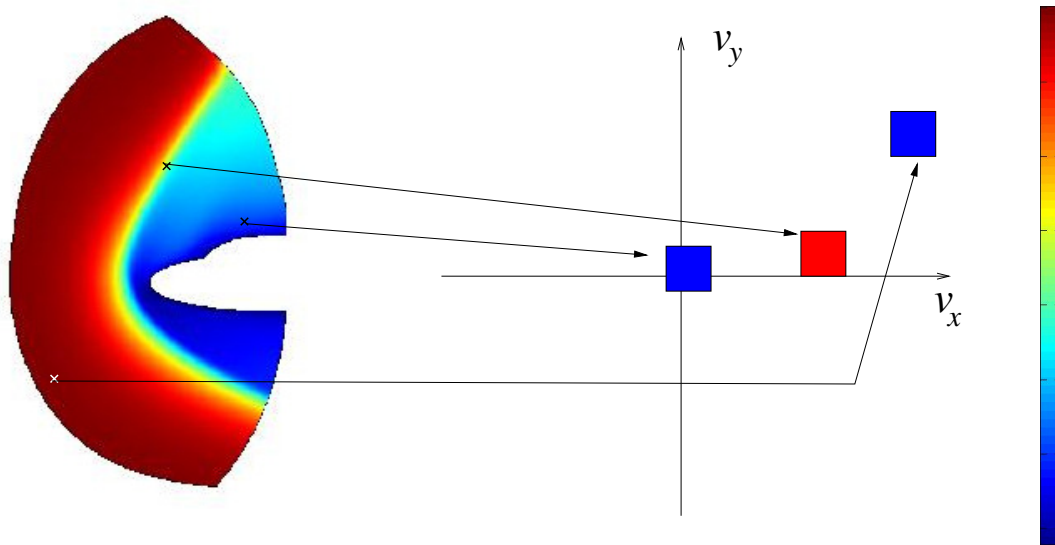
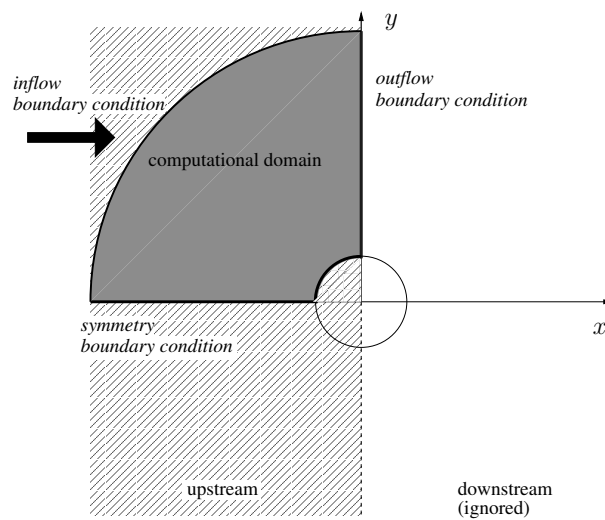


FIGURE 1.4: Three different distribution functions (top), the corresponding uniform Cartesian velocity grid (middle), and the grid locally refined in the support of the distributions (coarsened elsewhere).

FIGURE 1.5: The macroscopic CNS flow (left), and some values of  $\phi$  at three different  $v$  (right).FIGURE 1.6: Plane flow around a cylinder : geometry and computational domain. By symmetry with respect to the axis  $x$ , the computational domain is defined for  $y > 0$  only. The downstream flow is not simulated.

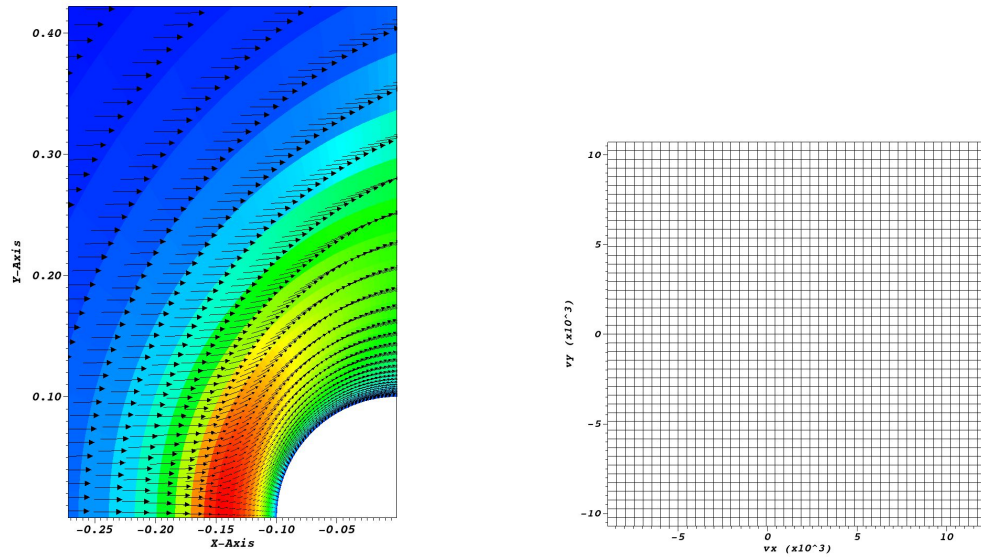


FIGURE 1.7: Plane flow around a cylinder : CNS velocity and temperature fields (left), corresponding fine Cartesian velocity grid (right).

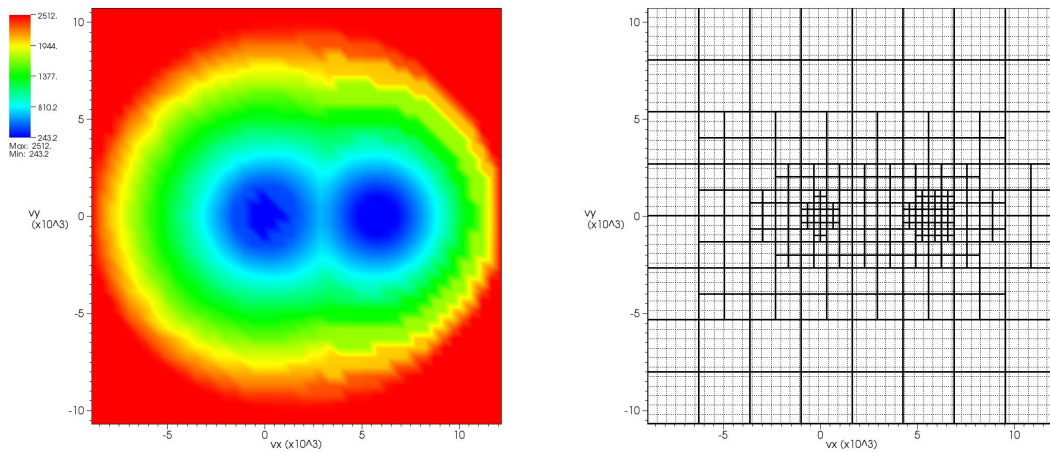


FIGURE 1.8: Plane flow around a cylinder. Left : support function, Right : velocity grids (solid line : induced AMR velocity grid, dotted line : initial fine Cartesian grid).

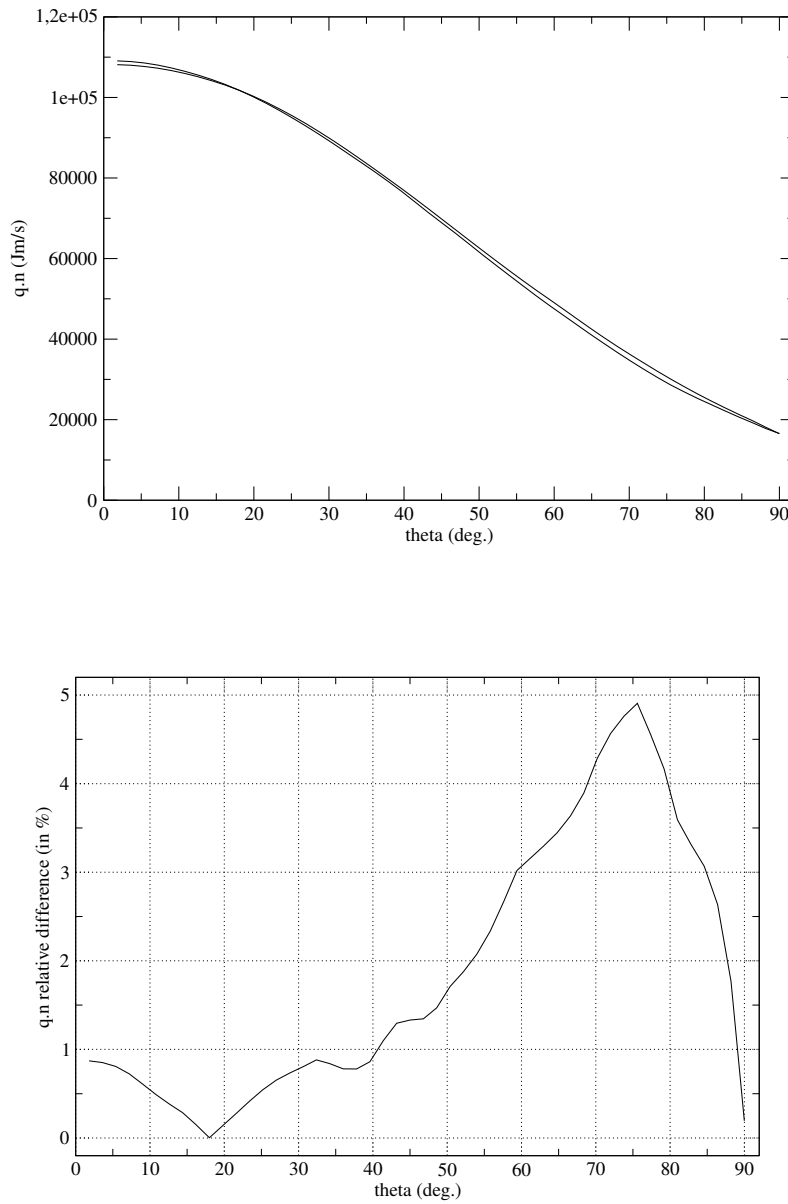


FIGURE 1.9: Plane flow around a cylinder. Comparison of the component of the heat flux normal to the solid wall : fluxes obtained with the fine Cartesian grid and the AMR grid (top), relative difference in percent (bottom). The angle  $\theta$  defines the position of the point along the solid boundary, such that  $\theta=0$  at the stagnation point and  $\theta=90$  at the end of the wall.

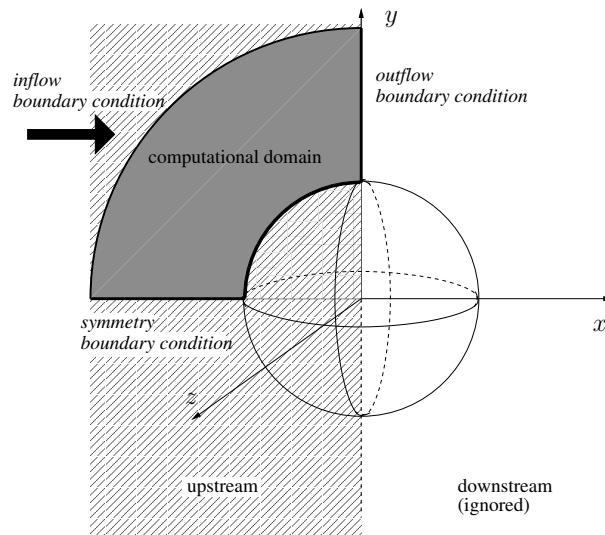


FIGURE 1.10: Axisymmetric flow around a sphere : geometry and computational domain. By rotational symmetry and by symmetry with respect to the axis  $x$ , the computational domain is included in the plane  $(x, y, 0)$ . The downstream flow is not simulated.

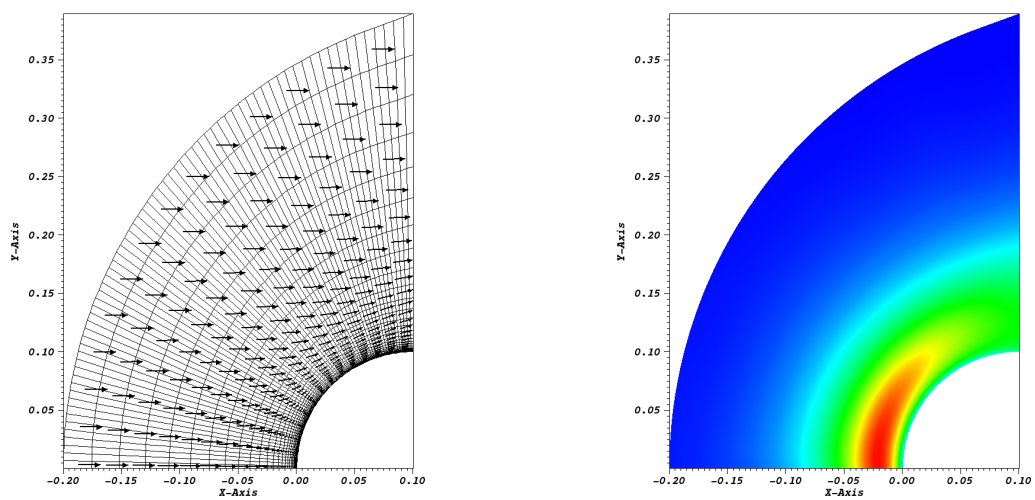


FIGURE 1.11: Axisymmetric flow around a sphere : velocity field, spatial mesh and temperature field.



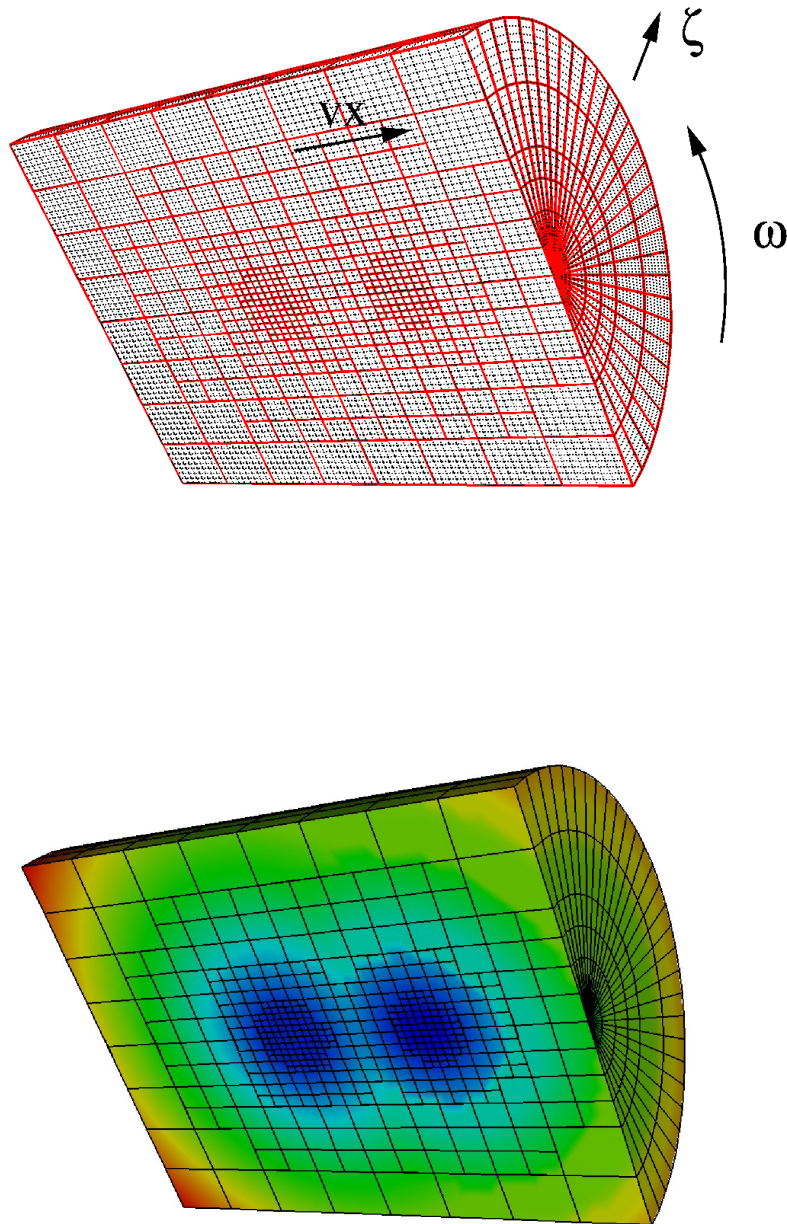


FIGURE 1.12: 2D axisymmetric flow around a sphere. Top : uniform cylindrical velocity grid (dashed lines) and AMR velocity grid (solid) with the cylindrical velocity coordinate system  $(v_x, \zeta, \omega)$ . Bottom : The same cylindrical AMR velocity grid with the contours of the support function based on the CNS pre-simulation.

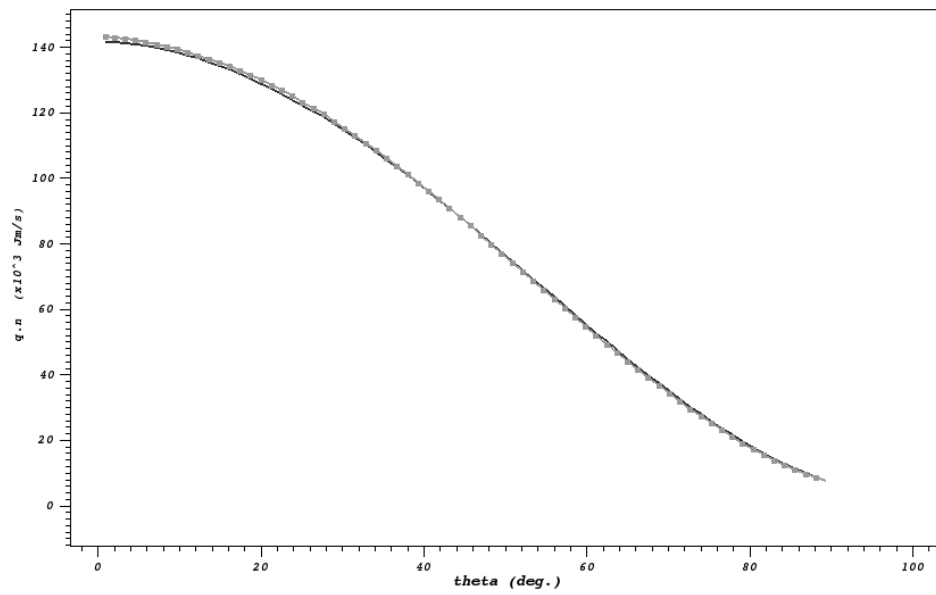


FIGURE 1.13: 2D axisymmetric flow around a sphere : comparison of the component of the heat flux normal to the surface : fluxes obtained with the fine uniform velocity grid (dot) and the AMR velocity grid (solid)

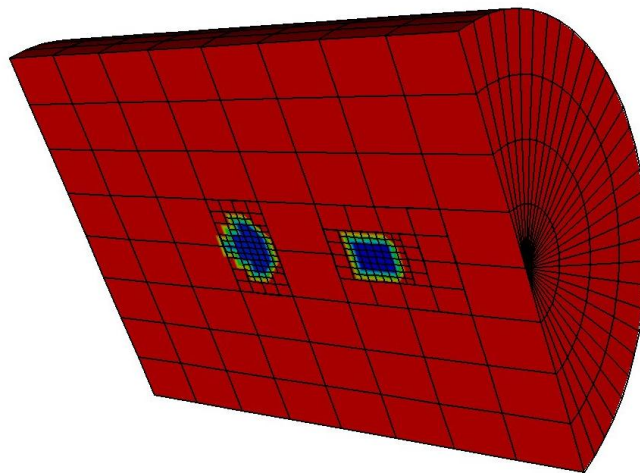


FIGURE 1.14: 2D axisymmetric flow around a sphere : cylindrical AMR velocity grid computed with Rankine-Hugoniot relations rather than with the CNS pre-simulation, with contours of the corresponding support function. Note that the refined zones are much smaller than in the previous grid, and that the contours of the support function are different (see figure 1.12)

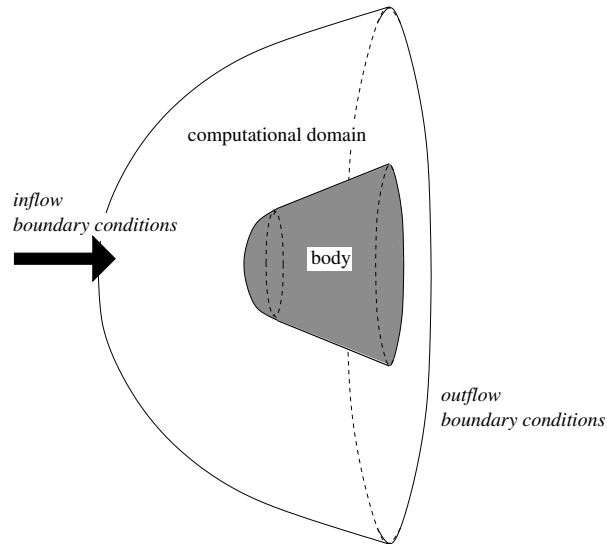


FIGURE 1.15: 3D flow around a cone : geometry and computational domain. The downstream flow is not simulated.

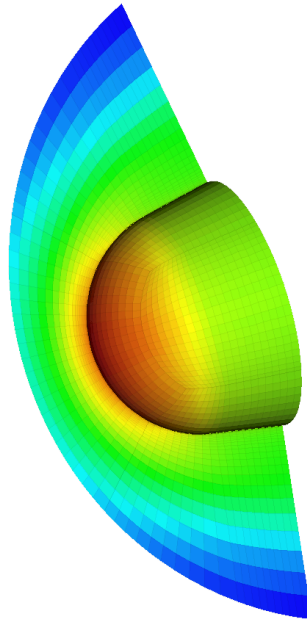


FIGURE 1.16: 3D flow around a cone : pressure field computed with the kinetic code and the AMR velocity grid. The pressure is shown on the surface of the solid, as well as in a vertical plane  $(x, y, 0)$  that crosses the cone along its middle part.

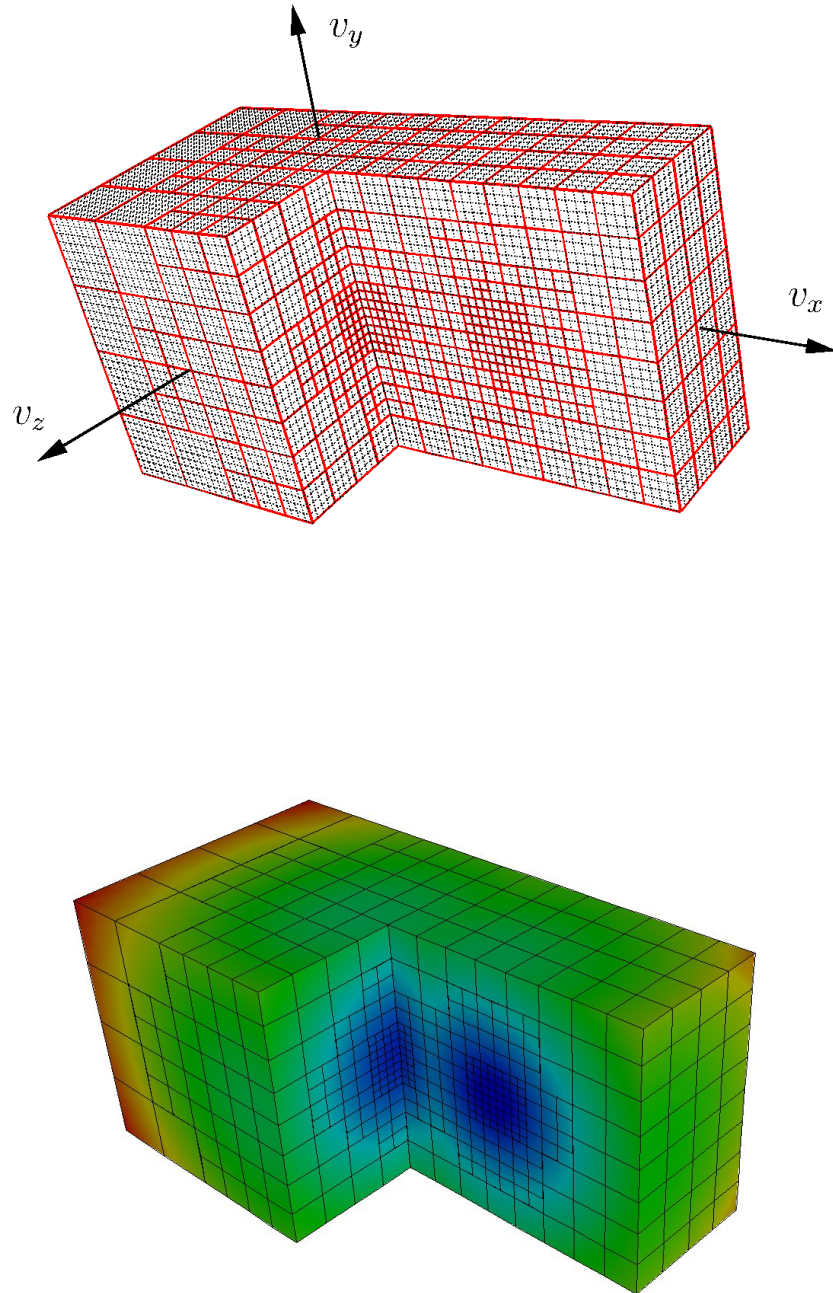


FIGURE 1.17: 3D flow around a cone. Top : the 3D Cartesian velocity grid (dashed line) and AMR velocity grid (solid) with the Cartesian coordinate system  $(v_x, v_y, v_z)$ . In order to see the refined zones around  $(0,0,0)$  and  $(u_{upstream}, 0, 0)$  that are inside the grid, a part of the grid has been hidden. Bottom : The same AMR velocity grid with the contours of the corresponding support function based on the CNS pre-simulation.

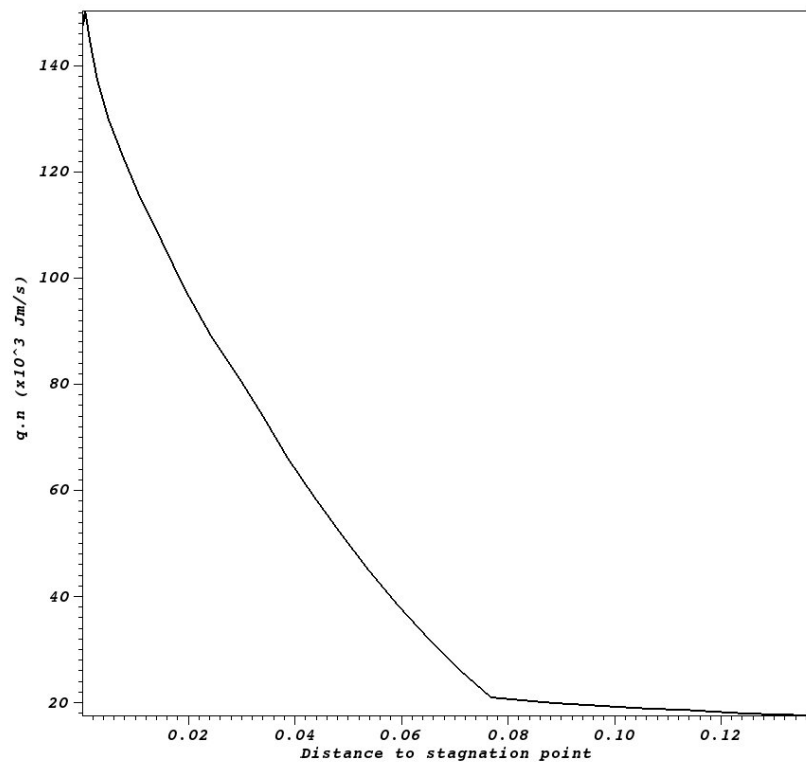


FIGURE 1.18: 3D flow around a cone : normal heat flux along the boundary (computed with the kinetic code and the AMR velocity grid).

## Chapitre 2

# Analyse asymptotique de l'équation BGK linéarisée avec condition de réflexion diffuse aux parois

*Chacun doit trouver sa solution.*

Woody Allen - *Maris et femmes*

La présence d'un corps solide au sein d'un écoulement raréfié entraîne l'apparition le long des parois du corps d'une couche limite cinétique, dite "couche de Knudsen". Cette couche limite se caractérise par une forte variation des grandeurs thermodynamiques (densité, vitesse, température, pression...) qui crée un écart entre ces grandeurs évaluées de part et d'autre de l'interface fluide/solide. On entre alors dans un régime fluide dit de "glissement", dans lequel les conditions aux limites classiques d'adhérence utilisées avec les équations d'Euler ou de Navier-Stokes ne sont plus valides.

De nombreuses formules, plus ou moins empiriques, ont été proposées pour quantifier ce glissement (voir par exemple la revue faite par Colin dans [28] pour le cas du transfert de chaleur convectif). On se propose dans cette partie d'étudier rigoureusement la limite fluide de l'équation BGK linéarisée avec des conditions aux limites de réflexion diffuse, d'établir les équations limites régissant les grandeurs thermodynamiques dans le cadre de cette linéarisation, puis les conditions aux limites associées en déterminant l'origine et les caractéristiques de la couche de Knudsen. Quelques auteurs ont déjà effectué cette analyse, notamment Yoshio Sone dans ses ouvrages de référence [63] et [64] ; cependant, il a paru nécessaire de clarifier ces travaux généraux peu accessibles afin de bien mettre en évidence la phénoménologie à l'oeuvre dans le cas qui nous intéresse.

On chemine en quatre étapes. On commence par adimensionner l'équation BGK stationnaire sans condition aux limites, introduisant ainsi le paramètre de raréfaction, le nombre de Knudsen. On linéarise ensuite l'opérateur BGK adimensionnel obtenu, en se plaçant dans une hypothèse de petites vitesses. Dans un troisième temps, on calcule la limite fluide de l'équation linéarisée en faisant tendre le nombre de Knudsen vers 0, ce qui permet d'établir le jeu d'équations de Stokes. Enfin, on introduit dans le calcul asymptotique la condition aux limites de réflexion diffuse pour aboutir à la notion de couche de Knudsen et aux conditions aux limites de glissement.

## 2.1 Adimensionnement de l'équation BGK stationnaire

L'équation BGK stationnaire s'écrit :

$$\mathbf{v} \cdot \nabla_x F(\mathbf{x}, \mathbf{v}) = Q(F)(\mathbf{x}, \mathbf{v}), \quad \mathbf{x} \in \Omega, \quad \mathbf{v} \in \mathbb{R}^3, \quad (2.1)$$

où le terme  $Q(F)$ , traduisant l'effet des collisions des particules de gaz, de fréquence moyenne  $\tau^{-1}$ , est modélisé par une relaxation de la distribution  $F$  vers l'équilibre thermodynamique associé  $\mathcal{M}(F)$  :

$$Q(F) = \frac{1}{\tau} (\mathcal{M}(F) - F). \quad (2.2)$$

Cet équilibre s'exprime en fonction de la densité  $\rho$ , de la vitesse macroscopique  $\mathbf{u}$  et de la température  $T$  sous forme d'une distribution maxwellienne des vitesses microscopiques  $\mathbf{v}$

$$\mathcal{M}(F) = M[\rho, \mathbf{u}, T] = \frac{\rho}{(2\pi RT)^{3/2}} \exp\left(-\frac{|\mathbf{u} - \mathbf{v}|^2}{2RT}\right),$$

où  $R$  est la constante spécifique du gaz.

On considère un état macroscopique de référence  $(\rho_0, \mathbf{u}_0, T_0)$ , et on suppose que le gaz est proche de l'état de repos, soit  $\mathbf{u}_0 \sim \mathbf{0}$ .

Pour adimensionner l'équation, on choisit deux longueurs de référence, une pour l'échelle macroscopique et une pour l'échelle microscopique, respectivement la longueur caractéristique du problème  $L$ , et le libre parcours moyen des molécules de gaz  $\lambda$ . On définit alors des grandeurs caractéristiques pour la position, la densité et la température :

$$\begin{cases} x^* = L, \\ \rho^* = \rho_0, \\ T^* = T_0. \end{cases}$$

La vitesse macroscopique étant supposée très faible, on choisit la vitesse caractéristique  $v^*$  au niveau microscopique (vitesse d'agitation thermique) :

$$v^* = \sqrt{RT_0} = \sqrt{RT^*}.$$

Les grandeurs adimensionnées correspondantes sont notées  $\hat{x}, \hat{\rho}, \hat{T}, \hat{\mathbf{v}}$ .

La maxwellienne s'écrit alors :

$$\begin{aligned} M[\rho, \mathbf{u}, T] &= \frac{\rho}{(2\pi RT)^{3/2}} \exp\left(-\frac{|\mathbf{u} - \mathbf{v}|^2}{2RT}\right) = \frac{\rho^* \hat{\rho}}{(2\pi RT^* \hat{T})^{3/2}} \exp\left(-\frac{v^{*2} |\hat{\mathbf{u}} - \hat{\mathbf{v}}|^2}{2RT^* \hat{T}}\right) \\ &= \frac{\rho^*}{v^{*3}} \frac{\hat{\rho}}{(2\pi \hat{T})^{3/2}} \exp\left(-\frac{|\hat{\mathbf{u}} - \hat{\mathbf{v}}|^2}{2\hat{T}}\right), \end{aligned}$$

on obtient ainsi la maxwellienne adimensionnée

$$\hat{M}[\hat{\rho}, \hat{\mathbf{u}}, \hat{T}] = \frac{\hat{\rho}}{(2\pi \hat{T})^{3/2}} \exp\left(-\frac{|\hat{\mathbf{u}} - \hat{\mathbf{v}}|^2}{2\hat{T}}\right),$$

et une valeur caractéristique pour la distribution  $F$  :

$$F^* = \frac{\rho^*}{v^{*3}}. \quad (2.3)$$

On adimensionne alors le terme de collision (2.2) :

$$Q(F) = \frac{F^*}{\tau^*} \frac{1}{\hat{\tau}} \left( \hat{M}[\hat{\rho}, \hat{\mathbf{u}}, \hat{T}] - \hat{F} \right). \quad (2.4)$$

Le premier membre de l'équation (2.1) s'adimensionne quant à lui de la façon suivante :

$$\mathbf{v} \cdot \nabla_x F = v^* \frac{F^*}{x^*} \hat{\mathbf{v}} \cdot \nabla_x \hat{F} = v^* \frac{F^*}{L} \hat{\mathbf{v}} \cdot \nabla_x \hat{F},$$

ce qui donne finalement :

$$\hat{\mathbf{v}} \cdot \nabla_x \hat{F} = \frac{L}{v^* \tau^*} \frac{1}{\hat{\tau}} \left( \hat{M}[\hat{\rho}, \hat{\mathbf{u}}, \hat{T}] - \hat{F} \right). \quad (2.5)$$

En utilisant la définition la plus courante du libre parcours moyen  $\lambda$  (plusieurs variantes existent)

$$\lambda = \tau \sqrt{RT},$$

ce qui donne

$$\tau = \frac{\lambda}{\sqrt{RT}} = \frac{\lambda^* \hat{\lambda}}{v^* \sqrt{\hat{T}}},$$

et donc finalement

$$\tau^* = \frac{\lambda^*}{v^*}, \quad \hat{\tau} = \frac{\hat{\lambda}}{\sqrt{\hat{T}}}. \quad (2.6)$$

Le temps de relaxation adimensionnel  $\hat{\tau}$  peut également s'exprimer en fonction des grandeurs thermodynamiques adimensionnelles  $\hat{\rho}$  et  $\hat{T}$ . En effet, le temps de relaxation peut s'écrire d'un point de vue macroscopique en fonction de la pression  $p$  et de la viscosité dynamique  $\mu$  :

$$\tau = \frac{\mu}{p}.$$

La loi d'état des gaz parfaits donne

$$p = \rho RT = \rho^* v^{*2} \hat{\rho} \hat{T} \Rightarrow \boxed{\begin{cases} p^* = \rho^* v^{*2} \\ \hat{p} = \hat{\rho} \hat{T} ; \end{cases}}$$

la viscosité dynamique  $\mu$  peut quant à elle s'écrire

$$\mu = \mu_{ref} \left( \frac{T}{T_{ref}} \right)^\omega = \mu_{ref} \left( \frac{T^*}{T_{ref}} \right)^\omega \hat{T}^\omega \Rightarrow \boxed{\begin{cases} \mu^* = \mu_{ref} \left( \frac{T^*}{T_{ref}} \right)^\omega \\ \hat{\mu} = \hat{T}^\omega, \end{cases}}$$

où  $\mu_{ref}$ ,  $T_{ref}$  et  $\omega$  sont des grandeurs fixées. Ceci donne finalement

$$\frac{1}{\hat{\tau}} = \frac{\hat{p}}{\hat{\mu}} = \hat{\rho} \hat{T}^{(1-\omega)}. \quad (2.7)$$



On peut alors réécrire l'équation (3.22) en utilisant l'expression de  $\tau^*$  obtenue avec (2.6) et celle de  $\hat{\tau}$  obtenue avec (2.7) :

$$\hat{\mathbf{v}} \cdot \nabla_x \hat{F} = \frac{L}{\lambda^*} \hat{\rho} \hat{T}^{(1-\omega)} (\hat{M}[\hat{\rho}, \hat{\mathbf{u}}, \hat{T}] - \hat{F}) = \frac{L}{\lambda^*} \hat{Q}(\hat{F}) = \frac{1}{\text{Kn}} \hat{Q}(\hat{F}). \quad (2.8)$$

La grandeur adimensionnelle  $\text{Kn} = \frac{\lambda^*}{L}$  est appelée nombre de Knudsen, caractérisant l'état de raréfaction de l'écoulement.

On peut alors exprimer les moments de la distribution adimensionnelle  $\hat{F}$  :

$$\left\{ \begin{array}{l} \rho = \int_{\mathbb{R}^3} F d\mathbf{v} \Rightarrow \rho^* \hat{\rho} = F^* v^{*3} \int_{\mathbb{R}^3} \hat{F} d\hat{\mathbf{v}} \Rightarrow \boxed{\hat{\rho} = \int_{\mathbb{R}^3} \hat{F} d\hat{\mathbf{v}},} \\ \rho \mathbf{u} = \int_{\mathbb{R}^3} \mathbf{v} F d\mathbf{v} \Rightarrow \rho^* v^* \hat{\rho} \hat{\mathbf{u}} = v^* F^* v^{*3} \int_{\mathbb{R}^3} \hat{\mathbf{v}} \hat{F} d\hat{\mathbf{v}} \Rightarrow \boxed{\hat{\rho} \hat{\mathbf{u}} = \int_{\mathbb{R}^3} \hat{\mathbf{v}} \hat{F} d\hat{\mathbf{v}},} \\ 3\rho RT = \int_{\mathbb{R}^3} |\mathbf{u} - \mathbf{v}|^2 F d\mathbf{v} \Rightarrow \rho^* v^{*2} 3\hat{\rho} \hat{T} = v^{*2} F^* v^{*3} \int_{\mathbb{R}^3} |\hat{\mathbf{u}} - \hat{\mathbf{v}}|^2 \hat{F} d\hat{\mathbf{v}} \\ \Rightarrow \boxed{3\hat{\rho} \hat{T} = \int_{\mathbb{R}^3} |\hat{\mathbf{u}} - \hat{\mathbf{v}}|^2 \hat{F} d\hat{\mathbf{v}},} \end{array} \right. \quad (2.9)$$

☞ NB : il est possible d'adimensionner autrement. En effet, comme évoqué plus haut, le choix de la vitesse dans l'adimensionnement de  $\lambda$  offre plusieurs possibilités. Un choix très classique (fait par exemple par Sone dans [64]) est d'écrire

$$\lambda = \tau \xi,$$

où  $\xi$  est la vitesse moyenne des molécules

$$\xi = \frac{1}{\rho} \int_{\mathbb{R}^3} |\mathbf{v}| F d\mathbf{v}, \quad |\mathbf{v}| = \|\mathbf{v}\|_{L^2} = (v_i v_i)^{1/2},$$

que l'on adimensionne avec (2.3)

$$\xi = \frac{v^{*4} F^*}{\rho^*} \frac{1}{\hat{\rho}} \int_{\mathbb{R}^3} |\hat{\mathbf{v}}| \hat{F} d\hat{\mathbf{v}} = v^* \frac{1}{\hat{\rho}} \int_{\mathbb{R}^3} |\hat{\mathbf{v}}| \hat{F} d\hat{\mathbf{v}} = \xi^* \hat{\xi},$$

soit

$$\xi^* = v^*,$$

et

$$\tau = \frac{\lambda}{\xi} = \frac{\lambda}{v^*} \hat{\rho} \left( \int_{\mathbb{R}^3} |\hat{\mathbf{v}}| \hat{F} d\hat{\mathbf{v}} \right)^{-1},$$

ce qui donne pour le temps de relaxation la même grandeur caractéristique que dans la relation (2.6), mais une grandeur adimensionnelle différente :

$$\tau^* = \frac{\lambda}{v^*}, \quad \hat{\tau} = \hat{\rho} \left( \int_{\mathbb{R}^3} |\hat{\mathbf{v}}| \hat{F} d\hat{\mathbf{v}} \right)^{-1}. \quad (2.10)$$

Ce choix d'adimensionnement est légitime, cependant il présente l'inconvénient de ne pas respecter l'invariance galiléenne, ce qui n'est pas le cas de l'adimensionnement utilisant la vitesse d'agitation thermique comme grandeur caractéristique.  $\square$

On procède maintenant à la linéarisation de l'équation BGK linéarisée (2.8) autour de l'état de référence choisi.

## 2.2 Linéarisation de l'équation adimensionnée

On suppose désormais que la distribution  $F$  est proche de l'équilibre associé à l'état de référence  $(\rho_0, \mathbf{u}_0, T_0)$  :

$$F \sim M[\rho_0, \mathbf{u}_0, T_0] \sim M[\rho_0, \mathbf{0}, T_0],$$

et on va procéder à la linéarisation du terme de relaxation autour de cet équilibre.

On commence par exprimer le problème avec les variables adimensionnelles en calculant la distribution maxwellienne adimensionnée associée à l'état de référence ; compte-tenu de l'adimensionnement choisi,

$$(\rho, \mathbf{u}, T) = (\rho_0, \mathbf{0}, T_0) \Rightarrow (\hat{\rho}, \hat{\mathbf{u}}, \hat{T}) = (1, \mathbf{0}, 1),$$

et donc

$$\hat{M}(\hat{\rho}, \hat{\mathbf{u}}, \hat{T}) = \hat{M}[1, \mathbf{0}, 1] = \frac{1}{(2\pi)^{3/2}} \exp\left(-\frac{|\hat{\mathbf{v}}|^2}{2}\right).$$

La distribution d'équilibre  $\hat{M}[1, \mathbf{0}, 1]$  est appelée "maxwellienne absolue". On va donc linéariser le terme de relaxation adimensionné (2.4) autour de cette maxwellienne, ceci de deux façons différentes : en exprimant les équilibres en fonction des distributions (section 2.2.1) ou en fonction des variables thermodynamiques (section 3.2) .

Dans toute la suite, on ne considère plus que les grandeurs adimensionnelles, que l'on note désormais comme les grandeurs dimensionnées pour simplifier les écritures.

### 2.2.1 Expression formelle : perturbation des fonctions de distribution

On considère l'équation adimensionnelle :

$$\mathbf{v} \cdot \nabla_x F = \frac{1}{\text{Kn}} Q(F) ; \quad (2.11)$$

on perturbe alors  $F$  à partir de son état d'équilibre (ce qui revient à une "décomposition micro-macro"), en introduisant une perturbation microscopique  $f$  :

$$F = M[1, \mathbf{0}, 1] + f, \quad f \ll 1. \quad (2.12)$$

En remplaçant dans l'équation (2.11), il vient :

$$\mathbf{v} \cdot \nabla_x (M[1, \mathbf{0}, 1]) + \mathbf{v} \cdot \nabla_x f = \frac{1}{\text{Kn}} Q(M[1, \mathbf{0}, 1] + f).$$

En linéarisant le terme de relaxation autour de  $M[1, \mathbf{0}, 1]$ , on obtient

$$\mathbf{v} \cdot \nabla_x (M[1, \mathbf{0}, 1]) + \mathbf{v} \cdot \nabla_x f = \frac{1}{\text{Kn}} \left[ Q(M[1, \mathbf{0}, 1]) + Q'(M[1, \mathbf{0}, 1]) \cdot f + \mathcal{O}(f^2) \right],$$

soit :

$$\mathbf{v} \cdot \nabla_x f = \frac{1}{\text{Kn}} \left[ Q'(M[1, \mathbf{0}, 1]) \cdot f + \mathcal{O}(f^2) \right].$$

L'équation BGK linéarisée s'écrit donc

$$\boxed{\mathbf{v} \cdot \nabla_x f = \frac{1}{\text{Kn}} L(f)}, \quad (2.13)$$

avec l'opérateur BGK linéarisé

$$L(f) = Q'(M[1, \mathbf{0}, 1]) \cdot f = \frac{\partial Q}{\partial F}(M[1, \mathbf{0}, 1]) \cdot f. \quad (2.14)$$

On a

$$Q(F) = \frac{1}{\tau(F)}(\mathcal{M}(F) - F) \Rightarrow \frac{\partial Q}{\partial F}(G) = \frac{1}{\tau(F)} \left( \frac{\partial \mathcal{M}}{\partial F}(G) - 1 \right) - \frac{1}{\tau^2(F)} \frac{\partial \tau}{\partial F}(G)(\mathcal{M}(G) - G),$$

ce qui donne, sachant  $\mathcal{M}(M[1, \mathbf{0}, 1]) = M[1, \mathbf{0}, 1]$  :

$$\frac{\partial Q}{\partial F}(M[1, \mathbf{0}, 1]) \cdot f = \frac{1}{\tau(M[1, \mathbf{0}, 1])} \left( \frac{\partial \mathcal{M}}{\partial F}(M[1, \mathbf{0}, 1]) \cdot f - f \right).$$

La relation (2.7) donne  $\tau(M[1, \mathbf{0}, 1]) = 1$ , soit

$$L(f) = \frac{\partial \mathcal{M}}{\partial F}(M[1, \mathbf{0}, 1]) \cdot f - f.$$

☞ NB : avec l'adimensionnement du temps de relaxation utilisant la vitesse moyenne des molécules,

$$F \sim M[1, \mathbf{0}, 1] \Rightarrow \tau = \left( \int_{\mathbb{R}^3} |\mathbf{v}| M[1, \mathbf{0}, 1] d\mathbf{v} \right)^{-1} = \frac{\sqrt{2\pi}}{2} \quad \text{selon (2.10),}$$

ce qui fait apparaître un facteur différent devant  $L(f)$ .  $\square$

L'expression formelle donnée ci-dessus étant difficilement calculable, on exprime l'opérateur  $L(f)$  en perturbant les grandeurs thermodynamiques.

### 2.2.2 Expression de l'opérateur BGK linéarisé en fonction des moments

On cherche à déterminer les perturbations  $(\rho', \mathbf{u}', T') = \mathcal{O}(f)$  telles que :

$$\mathcal{M}(M[1, \mathbf{0}, 1] + f) = M[1 + \rho', \mathbf{u}', 1 + T'],$$

de façon à obtenir :

$$\begin{aligned} Q(M[1, \mathbf{0}, 1] + f) &= \frac{1}{\tau(M[1, \mathbf{0}, 1] + f)} (\mathcal{M}(M[1, \mathbf{0}, 1] + f) - M[1, \mathbf{0}, 1] - f) \\ &= \frac{1}{\tau(M[1, \mathbf{0}, 1] + f)} (M[1 + \rho', \mathbf{u}', 1 + T'] - M[1, \mathbf{0}, 1] - f). \end{aligned}$$

et linéariser  $M[1 + \rho', \mathbf{u}', 1 + T']$  autour de  $M[1, \mathbf{0}, 1]$ . Pour un état  $(\rho, \mathbf{u}, T)$  donné, la maxwellienne perturbée s'écrit

$$M[\rho + \rho', \mathbf{u} + \mathbf{u}', T + T'] = \frac{\rho + \rho'}{(2\pi(T + T'))^{3/2}} \exp\left(-\frac{|\mathbf{u} + \mathbf{u}' - \mathbf{v}|^2}{2(T + T')}\right);$$

en supposant  $\rho' \ll \rho$ ,  $\mathbf{u}' \ll \mathbf{u}$ ,  $T' \ll T$ , un développement limité au premier ordre donne (cf Annexe B.1)

$$M[\rho + \rho', \mathbf{u} + \mathbf{u}', T + T'] = M[\rho, \mathbf{u}, T] \left[ 1 + \frac{\rho'}{\rho} + \frac{\mathbf{v} - \mathbf{u}}{T} \cdot \mathbf{u}' + \left( \frac{|\mathbf{v} - \mathbf{u}|^2}{2T^2} - \frac{3}{2T} \right) T' \right] + \mathcal{O}(\rho'^2, \mathbf{u}'^2, T'^2),$$

soit pour  $(\rho, \mathbf{u}, T) = (1, \mathbf{0}, 1)$

$$\begin{aligned} M[1 + \rho', \mathbf{u}', 1 + T'] &= M[1, \mathbf{0}, 1] + \underbrace{\left[ \rho' + \mathbf{u}' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T' \right]}_{= \Pi(\rho', \mathbf{u}', T')} M[1, \mathbf{0}, 1] + \mathcal{O}(\rho'^2, \mathbf{u}'^2, T'^2), \end{aligned} \quad (2.15)$$

et

$$\begin{aligned} M[1 + \rho', \mathbf{u}', 1 + T'] - M[1, \mathbf{0}, 1] - f &= \Pi(\rho', \mathbf{u}', T') - f + \mathcal{O}(\rho'^2, \mathbf{u}'^2, T'^2). \\ &= \Pi(\rho', \mathbf{u}', T') - f + \mathcal{O}(f^2). \end{aligned}$$

Ce terme étant d'ordre 1, on ne développe le facteur  $\frac{1}{\tau(M[1, \mathbf{0}, 1] + f)}$  qu'à l'ordre 0 :

$$\frac{1}{\tau(M[1, \mathbf{0}, 1] + f)} = \frac{1}{\tau(M[1, \mathbf{0}, 1])} + \mathcal{O}(f) = 1 + \mathcal{O}(f),$$

soit

$$Q(M[1, \mathbf{0}, 1] + f) = \Pi(\rho', \mathbf{u}', T') - f + \mathcal{O}(f^2).$$

On obtient donc l'expression de l'opérateur BGK linéarisé en fonction des perturbations  $(\rho', \mathbf{u}', T')$  :

$$L(f) = \left[ \rho' + \mathbf{u}' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T' \right] M[1, \mathbf{0}, 1] - f = \Pi(\rho', \mathbf{u}', T') - f. \quad (2.16)$$

Il reste à déterminer l'expression de  $\rho'$ ,  $\mathbf{u}'$  et  $T'$  en fonction de  $f$ , ce qui se fait avec un calcul de moments. Pour simplifier les écritures, on adopte dans toute la suite la notation suivante :

$$\langle \cdot \rangle = \int_{\mathbb{R}^3} \cdot \, d\mathbf{v}.$$

On a  $L(f) = \Pi(\rho', \mathbf{u}', T') - f$ , d'où en notant  $\mathbf{m}(\mathbf{v}) = (1, \mathbf{v}, |\mathbf{v}|^2)^T$

$$\left\langle \mathbf{m}(\mathbf{v}) L(f) \right\rangle = \left\langle \mathbf{m}(\mathbf{v}) \Pi(\rho', \mathbf{u}', T') \right\rangle - \left\langle \mathbf{m}(\mathbf{v}) f \right\rangle,$$

or selon (2.14),

$$\begin{aligned} \left\langle \mathbf{m}(\mathbf{v}) L(f) \right\rangle &= \left\langle \mathbf{m}(\mathbf{v}) Q'(M[1, \mathbf{0}, 1]) \cdot f \right\rangle = \left\langle \mathbf{m}(\mathbf{v}) \lim_{\varepsilon \rightarrow 0} \frac{Q(M[1, \mathbf{0}, 1] + \varepsilon f) - Q(M[1, \mathbf{0}, 1])}{\varepsilon f} f \right\rangle \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left\langle \mathbf{m}(\mathbf{v}) \underbrace{Q(M[1, \mathbf{0}, 1] + \varepsilon f)}_{= 0 \quad \forall \varepsilon \in \mathbb{R}} \right\rangle \\ &= 0, \end{aligned}$$

ce qui donne selon (2.9), en notant  $(\rho, \mathbf{u}, T)$  les grandeurs macroscopiques issues de  $f$  :

$$\left\langle \mathbf{m}(\mathbf{v}) \Pi(\rho', \mathbf{u}', T') \right\rangle = \left\langle \mathbf{m}(\mathbf{v}) f \right\rangle = (\rho, \rho \mathbf{u}, \rho \mathbf{u}^2 + 3\rho T)^T.$$

Un simple calcul nous donne (cf Annexe B.2) :

$$\left\langle \mathbf{m}(\mathbf{v}) \Pi(\rho', \mathbf{u}', T') \right\rangle = \left( \rho', \mathbf{u}', 3(\rho' + T') \right)^T,$$

soit

$$\begin{cases} \rho' = \rho = \langle f \rangle, \\ \mathbf{u}' = \rho \mathbf{u} = \langle \mathbf{v} f \rangle, \\ T' = \frac{1}{3}(\rho \mathbf{u}^2 + 3\rho T) - \rho = \left\langle \frac{|\mathbf{v}|^2 - 3}{3} f \right\rangle. \end{cases} \quad (2.17)$$

On obtient donc finalement l'expression de  $L(f)$  en fonction des moments de  $f$  :

$$\boxed{L(f) = \left[ \langle f \rangle + \langle \mathbf{v} f \rangle \cdot \mathbf{v} + \left\langle \frac{|\mathbf{v}|^2 - 3}{3} f \right\rangle \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \right] M[1, \mathbf{0}, 1] - f.} \quad (2.18)$$

☞ NB : en considérant l'espace  $E$  des fonctions de distribution, généralement  $L^2(\Omega \times \mathbb{R}^3)$  muni du produit scalaire

$$\langle f, g \rangle = \int_{\mathbb{R}^3} f g M[1, \mathbf{0}, 1] d\mathbf{v} = \left\langle f g M[1, \mathbf{0}, 1] \right\rangle,$$

(ce choix d'espace est justifié au chapitre 3), on remarque que  $\Pi(\rho', \mathbf{u}', T') = \Pi(f)$  est une projection de  $f$  sur le sous-espace  $V$  de  $E$  engendré par les invariants collisionnels, soit

$$V = \text{Vect}\left(1, \mathbf{v}, \frac{|\mathbf{v}|^2 - 3}{2}\right).$$

(Ces trois invariants forment une base orthogonale de  $V$  pour ce produit scalaire). Des considérations algébriques simples permettent de le comprendre facilement. Les moments  $\langle \mathbf{m}(\mathbf{v}) f \rangle$  sont

dans cette configuration les composantes du projeté de  $f$  sur  $V$  ; l'égalité

$$\langle \mathbf{m}(\mathbf{v})Q(f) \rangle = \langle \mathbf{m}(\mathbf{v})(M(f) - f) \rangle = \mathbf{0}$$

signifie que  $Q(f) \in V^\perp$ . En notant  $M_\perp(f)$  le projeté de  $M(f)$  sur  $V^\perp$ , on voit ainsi que la non linéarité de  $Q$ , c'est-à-dire de  $f \mapsto M(f)$ , vient de celle de l'application  $f \mapsto M_\perp(f)$ . Effectuer la linéarisation  $M(f) \rightarrow \Pi(f)$  revient donc à supprimer cette composante sur  $V^\perp$  :  $\Pi(f)$  se ramène alors à une projection sur  $V$ .  $\square$

A partir de l'expression (2.18) de l'opérateur BGK linéarisé, on procède dans la partie suivante au calcul de la limite asymptotique de l'équation (2.13).

## 2.3 Limite incompressible

On part de l'équation BGK linéarisée (2.13) et on suppose

$$\text{Kn} = \varepsilon \ll 1. \quad (2.19)$$

On effectue alors un développement de Grad-Hilbert par rapport à  $\varepsilon$  sur la distribution  $f$  :

$$f(\mathbf{x}, \mathbf{v}, \varepsilon) = f_0(\mathbf{x}, \mathbf{v}) + \varepsilon f_1(\mathbf{x}, \mathbf{v}) + \varepsilon^2 f_2(\mathbf{x}, \mathbf{v}) + \dots \quad (2.20)$$

On a alors en réinjectant dans (2.13) et en utilisant la linéarité de l'opérateur  $L$  :

$$\begin{aligned} \mathbf{v} \cdot \nabla_x f &= \frac{1}{\text{Kn}} L(f) \Leftrightarrow \mathbf{v} \cdot \nabla_x (f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \dots) = \frac{1}{\varepsilon} L(f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \dots) \\ &\Leftrightarrow \mathbf{v} \cdot \nabla_x f_0 + \varepsilon \mathbf{v} \cdot \nabla_x f_1 + \varepsilon^2 \mathbf{v} \cdot \nabla_x f_2 + \dots = \frac{1}{\varepsilon} (L(f_0) + \varepsilon L(f_1) + \varepsilon^2 L(f_2) + \dots) \\ &\Leftrightarrow \mathbf{v} \cdot \nabla_x f_0 + \varepsilon \mathbf{v} \cdot \nabla_x f_1 + \varepsilon^2 \mathbf{v} \cdot \nabla_x f_2 + \dots = \varepsilon^{-1} L(f_0) + \varepsilon^0 L(f_1) + \varepsilon L(f_2) + \dots \end{aligned}$$

On peut alors identifier terme à terme, selon les puissances de  $\varepsilon$  :

•  $\varepsilon^{-1}$  :  $L(f_0) = 0$ , soit  $f_0 \in \text{Ker}(L)$ . En notant  $(\rho'_0, \mathbf{u}_0', T'_0)$  les moments de  $f_0$  définis selon (2.17), la relation (2.18) donne

$$f_0 = \left[ \rho'_0 + \mathbf{u}_0' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_0 \right] M[1, \mathbf{0}, 1]. \quad (2.21)$$

•  $\varepsilon^0$  :  $\mathbf{v} \cdot \nabla_x f_0 = L(f_1)$ . Cette relation impose la condition de résolubilité  $\mathbf{v} \cdot \nabla_x f_0 \in \text{Im}(L)$ . Or selon (2.18), pour toute distribution  $\varphi$ ,  $\langle \mathbf{m}(\mathbf{v})L(\varphi) \rangle = \mathbf{0}$ , donc

$$\mathbf{v} \cdot \nabla_x f_0 \in \text{Im}(L) \Rightarrow \langle \mathbf{m}(\mathbf{v})\mathbf{v} \cdot \nabla_x f_0 \rangle = \langle \mathbf{m}(\mathbf{v})L(f_1) \rangle = \mathbf{0} ;$$

les quantités  $(\rho'_0, \mathbf{u}_0', T'_0)$  doivent ainsi vérifier les trois relations :

➤  $0 = \left\langle \mathbf{v} \cdot \nabla_x f_0 \right\rangle = \left\langle \nabla \cdot (\mathbf{v} f_0) \right\rangle = \nabla \cdot \left\langle \mathbf{v} f_0 \right\rangle = \nabla \cdot \mathbf{u}_0'$  ce qui donne une équation de continuité

$$\boxed{\nabla \cdot \mathbf{u}_0' = 0.} \quad (2.22)$$

➤  $0 = \left\langle \mathbf{v} (\mathbf{v} \cdot \nabla_x f_0) \right\rangle = \nabla \cdot \left\langle \mathbf{v} \otimes \mathbf{v} f_0 \right\rangle = \nabla \cdot \left[ (\rho_0' + T_0') \bar{I} \right] = \nabla (\rho_0' + T_0')$ , soit la relation de Boussinesq

$$\boxed{\nabla (\rho_0' + T_0') = \mathbf{0}.} \quad (2.23)$$

➤ la relation  $\left\langle |\mathbf{v}|^2 (\mathbf{v} \cdot \nabla_x f_0) \right\rangle = 0$  donne le même résultat que la première.

Sous ces conditions :  $\mathbf{v} \cdot \nabla_x f_0 = L(f_1) = \Pi(\rho_1', \mathbf{u}_1', T_1') - f_1$ , , soit :

$$\boxed{f_1 = -\mathbf{v} \cdot \nabla_x f_0 + \left[ \rho_1' + \mathbf{u}_1' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T_1' \right] M[1, \mathbf{0}, 1].}$$

•  $\varepsilon^1$  :  $\mathbf{v} \cdot \nabla_x f_1 = L(f_2)$ . De la même façon que précédemment, on obtient une condition de résolubilité :  $\mathbf{v} \cdot \nabla_x f_1 \in \text{Im}(L)$ , qui donne

$$\mathbf{v} \cdot \nabla_x f_1 \in \text{Im}(L) \Rightarrow \left\langle \mathbf{m}(\mathbf{v}) \mathbf{v} \cdot \nabla_x f_1 \right\rangle = \left\langle \mathbf{m}(\mathbf{v}) L(f_2) \right\rangle = \mathbf{0},$$

et on obtient trois nouvelles relations sur les quantités  $(\rho', \mathbf{u}', T')$  :

$$\begin{aligned} \text{➤ } 0 &= \left\langle \mathbf{v} \cdot \nabla_x f_1 \right\rangle = \left\langle \mathbf{v} \cdot \nabla_x (-\mathbf{v} \cdot \nabla_x f_0) \right\rangle + \left\langle \mathbf{v} \cdot \nabla_x \Pi(\rho_1', \mathbf{u}_1', T_1') \right\rangle \\ &= -3 \underbrace{\Delta (\rho_0' + T_0')}_{=0} + 3 \nabla \cdot \mathbf{u}_1', \text{ soit une équation de continuité} \end{aligned}$$

$$\boxed{\nabla \cdot \mathbf{u}_1' = 0 ;}$$

$$\begin{aligned} \text{➤ } 0 &= \left\langle \mathbf{v} (\mathbf{v} \cdot \nabla_x f_1) \right\rangle = \left\langle \mathbf{v} (\mathbf{v} \cdot \nabla_x (-\mathbf{v} \cdot \nabla_x f_0)) \right\rangle + \left\langle \mathbf{v} (\mathbf{v} \cdot \nabla_x \Pi(\rho_1', \mathbf{u}_1', T_1')) \right\rangle \\ &= -\Delta (\mathbf{u}_0') + \nabla (\rho_1' + T_1') \quad , \text{ ce qui donne une équation de Stokes} \end{aligned}$$

$$\boxed{\nabla (\rho_1' + T_1') = \Delta (\mathbf{u}_0') ;}$$

➤ la dernière relation donne une équation de Laplace :

$$\boxed{\Delta (T_0') = 0.}$$

On a de la même façon que précédemment sous ces conditions l'expression de  $f_2$  :

$$f_2 = -\mathbf{v} \cdot \nabla_x f_1 + \left[ \rho'_2 + \mathbf{u}'_2 \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_2 \right] M[1, \mathbf{0}, 1].$$

En itérant le processus aux ordres supérieurs du développement de  $f$ , on obtient

$$\mathbf{v} \cdot \nabla_x f_n = L(f_{n+1}) \quad \forall n > 1,$$

ce qui donne l'ensemble des équations de Stokes :

$$\left\{ \begin{array}{l} \nabla \cdot \mathbf{u}'_n = 0, \\ \nabla (\rho'_n + T'_n) = \gamma_n \Delta (\mathbf{u}_{n-1}'), \\ \Delta (T'_{n-1}) = 0, \\ f_{n+1} = -\mathbf{v} \cdot \nabla_x f_n + \left[ \rho'_{n+1} + \mathbf{u}'_{n+1} \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_{n+1} \right] M[1, \mathbf{0}, 1]. \end{array} \right. \quad (2.24)$$

D'après [64], le coefficient  $\gamma_n$  est indépendant de  $n$  pour le modèle BGK, et s'exprime

$$\gamma_n = \frac{16}{15\sqrt{\pi}} \int_0^{+\infty} \zeta^6 e^{-\zeta^2} d\zeta = 1 \quad \forall n \geq 1 ;$$

nous ne prouvons pas ici cette expression, que nous avons pu vérifier pour  $n = 1$ , et que nous admettons pour les ordres supérieurs.

Par linéarité, on peut appliquer le développement de Hilbert sur les quantités  $(\rho', \mathbf{u}', T')$  :

$$\left\{ \begin{array}{l} \rho' = \langle f \rangle = \sum_n \varepsilon^n \langle f^n \rangle = \sum_n \varepsilon^n \rho'_n, \\ \mathbf{u}' = \langle \mathbf{v} f \rangle = \sum_n \varepsilon^n \langle \mathbf{v} f^n \rangle = \sum_n \varepsilon^n \mathbf{u}'_n, \\ T' = \left\langle \frac{|\mathbf{v}|^2 - 3}{3} f \right\rangle = \sum_n \varepsilon^n \left\langle \frac{|\mathbf{v}|^2 - 3}{3} f^n \right\rangle = \sum_n \varepsilon^n T'_n. \end{array} \right.$$

On a alors avec les relations (2.24) :

$$\left\{ \begin{array}{l} \nabla \cdot \mathbf{u}' = \sum_n \varepsilon^n \nabla \cdot \mathbf{u}'_n = 0, \\ \nabla (\rho' + T') = \nabla (\rho'_0 + T'_0) + \sum_{n>0} \varepsilon^n \nabla (\rho'_n + T'_n) = \sum_{n>0} \varepsilon^n \gamma_n \Delta (\mathbf{u}_{n-1}') = \varepsilon \Delta \mathbf{u}', \\ \Delta T' = \sum_n \varepsilon^n \Delta T'_n = 0. \end{array} \right.$$

La perturbation de l'écoulement est donc incompressible, vérifie la relation de Boussinesq et l'équation de Stokes avec un terme diffusif d'ordre 1 par rapport à  $\varepsilon$ .



On va maintenant considérer le problème BGK complet en introduisant la condition aux limites de réflexion diffuse utilisée au niveau des parois solides. On procède comme pour l'équation seule en trois étapes, adimensionnement, linéarisation et passage à la limite asymptotique.

## 2.4 Prise en compte des conditions aux limites

### 2.4.1 Expression adimensionnelle des conditions aux limites aux parois

On reprend dans cette partie les notations de la partie 2.1.

On définit sur la paroi  $\Gamma$  une normale **entrante** au domaine fluide  $\mathbf{n}(\mathbf{x})$  et deux tangentes  $\mathbf{t}(\mathbf{x})$  et  $\mathbf{s}(\mathbf{x})$ , telles que  $(\mathbf{n}, \mathbf{t}, \mathbf{s})$  soit un repère orthonormal direct :

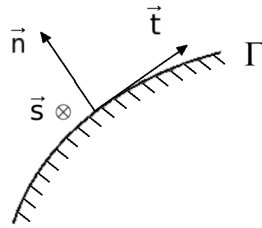


FIGURE 2.1: Vecteurs normal et tangentiels à la paroi

La condition de réflexion diffuse au niveau des parois solides s'écrit dans le cas dimensionnel :

$$F(\mathbf{x} \in \Gamma, \mathbf{v}) = \sigma(F)M_w(\mathbf{x}, \mathbf{v}) \text{ si } \mathbf{v} \cdot \mathbf{n}(\mathbf{x}) > 0, \quad M_w = M[1, \mathbf{u}_w(\mathbf{x}), T_w(\mathbf{x})],$$

où  $\mathbf{u}_w$  et  $T_w$  sont respectivement la vitesse et la température de la paroi, et où le coefficient de réflexion  $\sigma(F)$  s'écrit, en supposant que la vitesse de paroi est uniquement tangentielle ( $\mathbf{u}_w \cdot \mathbf{n} = 0$ ) :

$$\sigma(F) = -\frac{\int_{\mathbb{R}^3} (\mathbf{v} \cdot \mathbf{n}(\mathbf{x}))^- F(\mathbf{x}, \mathbf{v}) d\mathbf{v}}{\int_{\mathbb{R}^3} (\mathbf{v} \cdot \mathbf{n}(\mathbf{x}))^+ M_w d\mathbf{v}} = -\sqrt{\frac{2\pi}{RT_w}} \int_{\mathbb{R}^3} (\mathbf{v} \cdot \mathbf{n}(\mathbf{x}))^- F(\mathbf{x}, \mathbf{v}) d\mathbf{v}.$$

En adimensionnant comme précédemment dans la partie 2.1, on obtient

$$F(\mathbf{x} \in \Gamma, \mathbf{v}) = F^* \left( -\sqrt{\frac{2\pi}{\hat{T}_w}} \int_{\mathbb{R}^3} (\hat{\mathbf{v}} \cdot \mathbf{n})^- \hat{F} d\hat{\mathbf{v}} \right) \underbrace{\left( \frac{1}{(2\pi\hat{T}_w)^{3/2}} \exp\left(-\frac{|\hat{\mathbf{u}}_w - \hat{\mathbf{v}}|^2}{2\hat{T}_w}\right) \right)}_{= \hat{M}[1, \hat{\mathbf{u}}_w, \hat{T}_w]},$$

soit

$$\hat{F}|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = \hat{\sigma}(\hat{F}) \hat{M}[1, \hat{\mathbf{u}}_w, \hat{T}_w] = \hat{M}[\hat{\sigma}(\hat{F}), \hat{\mathbf{u}}_w, \hat{T}_w] \text{ si } \hat{\mathbf{v}} \cdot \mathbf{n} > 0, \quad (2.25)$$

avec

$$\hat{\sigma}(\hat{F}) = -\sqrt{\frac{2\pi}{\hat{T}_w}} \int_{\mathbb{R}^3} (\hat{\mathbf{v}} \cdot \mathbf{n})^- \hat{F} d\hat{\mathbf{v}} = -\frac{\int_{\mathbb{R}^3} (\hat{\mathbf{v}} \cdot \mathbf{n})^- \hat{F} d\hat{\mathbf{v}}}{\int_{\mathbb{R}^3} (\hat{\mathbf{v}} \cdot \mathbf{n})^+ \hat{M}_w d\hat{\mathbf{v}}}. \quad (2.26)$$

On reprend dans la suite les mêmes notations que précédemment.

La condition aux limites de réflexion diffuse est linéaire par rapport à  $\hat{F}$ . Cependant, la linéarisation de l'équation BGK repose sur l'hypothèse (2.12) que les distributions se ramènent à de petites perturbations de l'équilibre  $M[1, \mathbf{0}, 1]$ . On ne peut donc pas introduire cette condition aux limites telle quelle dans le problème linéarisé, on doit également la linéariser autour de la maxwellienne absolue.

## 2.4.2 Conditions aux limites et limite incompressible : couche limite cinétique

### 2.4.2.1 Problème BGK linéarisé avec conditions aux limites

La perturbation de l'écoulement autour de la distribution d'équilibre  $M[1, \mathbf{0}, 1]$  étudiée dans les paragraphes 2.2 et 2.3 vient nécessairement des parois. On considère donc le cas où la distribution d'équilibre induite par la paroi  $M[1, \mathbf{u}_w, T_w]$  est proche de  $M[1, \mathbf{0}, 1]$ , soit :

$$\begin{cases} \|\mathbf{u}_w\| \ll 1, \\ T_w = 1 + \theta_w, \quad \theta_w \ll 1. \end{cases}$$

Par linéarité du problème, on suppose que les perturbations des paramètres de paroi et la perturbation  $f$  qui en résulte dans la fonction de distribution sont du même ordre de grandeur :

$$(\|\mathbf{u}_w\|, \theta_w) = \mathcal{O}(f).$$

L'hypothèse de petites perturbations (2.12) introduite dans la condition aux limites (2.25) donne

$$\begin{aligned} F|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} &= (M[1, \mathbf{0}, 1] + f)|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = M[\sigma(F), \mathbf{u}_w, T_w] \\ &= M[1 + (\sigma(F) - 1), \mathbf{u}_w, T_w]. \end{aligned} \quad (2.27)$$

Le coefficient de réflexion  $\sigma(F)$  s'écrit selon la relation (2.26)

$$\begin{aligned} \sigma(F) &= \sigma(M[1, \mathbf{0}, 1] + f) = -\sqrt{\frac{2\pi}{T_w}} \left\langle (\mathbf{v} \cdot \mathbf{n})^- (M[1, \mathbf{0}, 1] + f) \right\rangle \\ &= -\sqrt{\frac{2\pi}{1 + \theta_w}} \left[ \left\langle (\mathbf{v} \cdot \mathbf{n})^- M[1, \mathbf{0}, 1] \right\rangle + \left\langle (\mathbf{v} \cdot \mathbf{n})^- f \right\rangle \right]. \end{aligned}$$

Or

$$\begin{cases} \frac{1}{\sqrt{1 + \theta_w}} = 1 - \frac{1}{2}\theta_w + \mathcal{O}(\theta_w^2) = 1 - \frac{1}{2}\theta_w + \mathcal{O}(f^2), \\ \left\langle (\mathbf{v} \cdot \mathbf{n})^- M[1, \mathbf{0}, 1] \right\rangle = \frac{-1}{\sqrt{2\pi}}, \\ \left\langle (\mathbf{v} \cdot \mathbf{n})^- f \right\rangle = \mathcal{O}(f), \end{cases}$$

et donc finalement :

$$\sigma(F) - 1 = -\frac{1}{2}\theta_w - \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f \right\rangle + \mathcal{O}(f^2).$$

On peut donc linéariser (2.27) selon (2.15) :

$$\begin{aligned} F|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} &= M[\sigma(F), \mathbf{u}_w, T_w] \approx M[1, \mathbf{0}, 1] + \Pi((\sigma(F) - 1), \mathbf{u}_w, \theta_w) \\ &= M[1, \mathbf{0}, 1] + \left[ (\sigma(F) - 1) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \right] M[1, \mathbf{0}, 1] \\ &= \left[ \sigma(F) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \right] M[1, \mathbf{0}, 1]. \end{aligned}$$

On obtient alors la condition aux limites linéarisée sur la perturbation  $f$  :

$$f|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = F|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} - M[1, \mathbf{0}, 1] = \left[ \sigma(F) - 1 + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \right] M[1, \mathbf{0}, 1],$$

soit

$$\boxed{\begin{cases} f|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = \left[ \sigma_w(f) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \right] M[1, \mathbf{0}, 1], \\ \sigma_w(f) = -\frac{1}{2}\theta_w - \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f \right\rangle. \end{cases}} \quad (2.28)$$

Le problème BGK linéarisé s'écrit donc :

$$\boxed{\begin{cases} \mathbf{v} \cdot \nabla_x f = \frac{1}{\text{Kn}} L(f), \\ f|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = \left[ \sigma_w(f) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \right] M[1, \mathbf{0}, 1], \\ \sigma_w(f) = -\frac{1}{2}\theta_w - \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f \right\rangle. \end{cases}}$$

On notera dans la suite :  $p_w(f) = \sigma_w(f) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w$ .

Le problème linéarisé complet étant déterminé, on peut maintenant analyser sa limite asymptotique comme il a déjà été fait sur l'équation seule au paragraphe 2.3.

#### 2.4.2.2 Développement de Grad-Hilbert : notion de couche limite cinétique

Le cas de l'équation linéarisée ayant déjà été traité, on ne s'intéresse ici qu'à la condition aux limites (2.28).

Le développement (2.20) introduit dans le coefficient de réflexion  $\sigma_w(f)$  donne :

$$\begin{aligned}\sigma_w(f) &= -\frac{1}{2}\theta_w - \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- (f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \dots) \right\rangle \\ &= -\frac{1}{2}\theta_w - \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_0 \right\rangle - \varepsilon \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_1 \right\rangle - \varepsilon^2 \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_2 \right\rangle \dots \\ &= \sigma_w(f_0) + \varepsilon \sigma_w(f_1) + \varepsilon^2 \sigma_w(f_2) + \dots\end{aligned}$$

où le coefficient  $\sigma_w$  se définit à chaque ordre :

$$\begin{cases} \sigma_w(f_0) = -\frac{1}{2}\theta_w - \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_0 \right\rangle, \\ \sigma_w(f_n) = -\sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_n \right\rangle \quad \forall n \in \mathbb{N}^*, \end{cases}$$

ce qui donne dans l'expression de  $p_w(f)$  :

$$\begin{aligned}p_w(f) &= \sigma_w(f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \dots) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \\ &= \sigma_w(f_0) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w + \varepsilon \sigma_w(f_1) + \varepsilon^2 \sigma_w(f_2) + \dots \\ &= p_w(f_0) + \varepsilon p_w(f_1) + \varepsilon^2 p_w(f_2) + \dots\end{aligned}$$

et donc également une définition à chaque ordre de  $p_w$  :

$$\begin{cases} p_w(f_0) = \sigma_w(f_0) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w, \\ p_w(f_n) = \sigma_w(f_n) \quad \forall n \in \mathbb{N}^*. \end{cases}$$

On obtient donc la décomposition suivante du problème BGK linéarisé :

$$\begin{cases} \mathbf{v} \cdot \nabla_x f_0 + \varepsilon \mathbf{v} \cdot \nabla_x f_1 + \varepsilon^2 \mathbf{v} \cdot \nabla_x f_2 + \dots = \varepsilon^{-1} L(f_0) + \varepsilon^0 L(f_1) + \varepsilon L(f_2) + \dots \\ f_0|_{\Gamma} + \varepsilon f_1|_{\Gamma} + \varepsilon^2 f_2|_{\Gamma} + \dots = \left[ p_w(f_0) + \varepsilon p_w(f_1) + \varepsilon^2 p_w(f_2) + \dots \right] M[1, \mathbf{0}, 1]. \end{cases} \quad (2.29)$$

La première équation de (2.29) donne, comme on l'a vu au paragraphe 2.3 en identifiant terme à terme les puissances de  $\varepsilon$ , le système

$$\begin{cases} f_0 = \left[ \rho'_0 + \mathbf{u}_0' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_0 \right] M[1, \mathbf{0}, 1], \\ \vdots \\ f_{n+1} = -(\mathbf{v} \cdot \nabla_x f_n) + \left[ \rho'_{n+1} + \mathbf{u}_{n+1}' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_{n+1} \right] M[1, \mathbf{0}, 1] \quad \forall n \in \mathbb{N}. \end{cases} \quad (2.30)$$

avec  $(\rho'_n, \mathbf{u}'_n, T'_n)$  les moments de  $f_n$  définis par la relation (2.17). La seconde équation donne par le même processus

$$\begin{cases} f_{0|\Gamma} = p_w(f_0) \quad M[1, \mathbf{0}, 1] = \left[ \sigma_w(f_0) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \right] M[1, \mathbf{0}, 1], \\ \vdots \\ f_{n+1|\Gamma} = p_w(f_{n+1}) \quad M[1, \mathbf{0}, 1] = \sigma_w(f_{n+1}) M[1, \mathbf{0}, 1] \quad \forall n \in \mathbb{N}. \end{cases} \quad (2.31)$$

Les systèmes (2.30) et (2.31) donnent respectivement à l'ordre 0

$$\begin{cases} f_0 = \left[ \rho'_0 + \mathbf{u}'_0 \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_0 \right] M[1, \mathbf{0}, 1], \\ f_{0|\Gamma} = \left[ \sigma_w(f_0) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \right] M[1, \mathbf{0}, 1]. \end{cases}$$

Il y a compatibilité des deux expressions avec les conditions d'adhérence

$$\boxed{\begin{cases} \rho'_0 = \sigma_w(f_0), \\ \mathbf{u}'_0 = \mathbf{u}_w, \\ T'_0 = \theta_w. \end{cases}} \quad (2.32)$$

Pour  $n \geq 1$ , il y a incompatibilité entre les expressions (2.30) et (2.31). En effet, le système (2.31) indique que sur le bord, chaque composante  $f_n$  s'exprime sous forme d'une maxwellienne par rapport à  $\mathbf{v}$ ; mais le terme contenant un gradient dans (2.30), reliant chaque ordre au précédent, introduit une pondération polynomiale.

On doit donc introduire dans le développement de Grad-Hilbert (2.20) un terme supplémentaire, actif au niveau des bords, qui permet de lever cette incompatibilité. La zone d'activité de ce terme est appelée couche limite cinétique (ou couche de Knudsen), on le nomme donc "correcteur de couche limite". La partie suivante est dédiée à l'étude de ce correcteur.

## 2.4.3 Correcteur de couche limite

### 2.4.3.1 Expression du correcteur

On suppose que la perturbation  $f$  possède deux composantes; la première, appelée "partie fluide", correspond à ce qui a été étudié précédemment, c'est-à-dire à la solution en l'absence de condition aux limites; on la note  $\bar{f}$ . La seconde composante est le correcteur, noté  $\varphi$ , qui permet de lever l'incompatibilité entre les systèmes (2.30) et (2.31) au niveau des bords solides. On écrit donc

$$f = \bar{f} + \varphi.$$

On a alors en introduisant cette décomposition dans l'équation (2.13)

$$\mathbf{v} \cdot \nabla_x f = \frac{1}{\varepsilon} L(f) \Rightarrow \mathbf{v} \cdot \nabla_x \bar{f} + \mathbf{v} \cdot \nabla_x \varphi = \frac{1}{\varepsilon} L(\bar{f}) + \frac{1}{\varepsilon} L(\varphi),$$

or  $\mathbf{v} \cdot \nabla_x \bar{f} = \frac{1}{\varepsilon} L(\bar{f})$ , ce qui permet d'obtenir l'équation vérifiée par le correcteur :

$$\boxed{\mathbf{v} \cdot \nabla_x \varphi = \frac{1}{\varepsilon} L(\varphi).} \quad (2.33)$$

On procède ensuite à l'étude de la limite asymptotique de la même façon que pour la partie fluide, en appliquant un développement de Hilbert sur le correcteur :

$$\varphi(\varepsilon) = \varphi_0 + \varepsilon \varphi_1 + \varepsilon^2 \varphi_2 + \dots \quad (2.34)$$

On a vu qu'il n'y avait pas de correction nécessaire à l'ordre 0, on a donc

$$\boxed{\varphi_0 = 0.} \quad (2.35)$$

On s'intéresse dans la suite au comportement du correcteur aux ordres suivants, en commençant par estimer la largeur de son support pour définir les variables pertinentes pour son analyse.

### 2.4.3.2 Variables de couche limite et problèmes de demi-espace

Une estimation simple d'ordre de grandeurs permet d'évaluer l'épaisseur de la couche limite cinétique. On a près du bord du domaine :

$$\mathbf{v} \cdot \nabla_x f_0 + \varepsilon \mathbf{v} \cdot \nabla_x f_1 = \frac{1}{\varepsilon} L(f_0) + L(f_1) + \varepsilon L(f_2) + \mathcal{O}(\varepsilon^2). \quad (2.36)$$

Compte-tenu de la relation (2.35),  $L(f_0) = L(\bar{f}_0) = 0$ .

On peut donc réécrire (2.36) :

$$\begin{aligned} \mathbf{v} \cdot \nabla_x f_0 - L(f_1) &= \varepsilon (L(f_2) - \mathbf{v} \cdot \nabla_x f_1) + \mathcal{O}(\varepsilon^2) \\ &= \varepsilon \underbrace{(L(f_2) - \mathbf{v} \cdot \nabla_x \bar{f}_1)}_{=0} - \mathbf{v} \cdot \nabla_x \varphi_1 + \mathcal{O}(\varepsilon^2) \\ &= -\varepsilon \mathbf{v} \cdot \nabla_x \varphi_1 + \mathcal{O}(\varepsilon^2). \end{aligned} \quad (2.37)$$

Or on a vu que la condition aux limites mettait l'expression de  $f_1$  donnée par le système (2.30) en défaut près de la paroi, soit

$$\mathbf{v} \cdot \nabla_x f_0 - L(f_1) \neq 0.$$

Cette différence localisée sur le bord subsiste si  $\varepsilon \rightarrow 0$  (même si la couche limite rétrécit du fait de l'approche du régime fluide). Elle est donc indépendante de  $\varepsilon$ , et on a donc

$$\mathbf{v} \cdot \nabla_x f_0 - L(f_1) = \mathcal{O}(1).$$

La relation (2.37) donne alors

$$\mathbf{v} \cdot \nabla_x \varphi_1 = \mathcal{O}(\varepsilon^{-1}).$$

Le correcteur de couche limite agit donc sur une épaisseur de l'ordre du nombre de Knudsen (soit du libre parcours moyen) dans la direction normale à la paroi, suivant le vecteur  $\mathbf{n}$  du repère local  $(\mathbf{n}, \mathbf{t}, \mathbf{s})$  défini sur la figure 2.1.

Il est alors judicieux de procéder à un redimensionnement suivant cette direction normale, en introduisant une nouvelle variable  $\eta$  quantifiant la distance à la paroi. La paroi est quant à elle paramétrée par les variables  $(\tau, \chi) : \Gamma = \Gamma(\tau, \chi)$ .

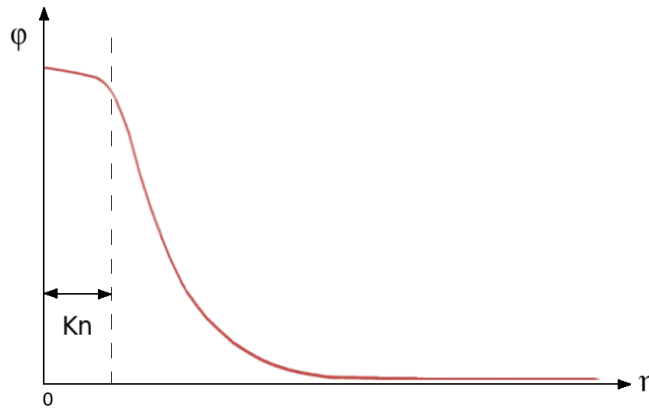


FIGURE 2.2: Allure du correcteur de couche limite en fonction de la distance à la paroi

On définit la variable  $\eta$  en posant :  $\mathbf{x} = (x, y, z) = \varepsilon \eta \mathbf{n}(\tau, \chi) + \mathbf{x}_\Gamma$ ,  $\mathbf{x}_\Gamma = (\tau, \chi)^T$ , soit :

$$\eta = \frac{1}{\varepsilon} \mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_\Gamma). \quad (2.38)$$

On exprime alors le premier membre de l'équation (2.33) avec ces nouvelles variables. En posant par commodité d'écriture  $(x, y, z) = (x_1, x_2, x_3)$  et  $(v_x, v_y, v_z) = (v_1, v_2, v_3)$  :

$$\begin{aligned} \mathbf{v} \cdot \nabla_x \varphi &= v_i \frac{\partial \varphi}{\partial x_i} = v_i \left( \frac{\partial \varphi}{\partial \eta} \frac{\partial \eta}{\partial x_i} + \frac{\partial \varphi}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right) \\ &= v_i \left( \frac{\partial \varphi}{\partial \eta} \frac{1}{\varepsilon} \mathbf{n} \cdot \mathbf{e}_i + \frac{\partial \varphi}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right) \\ &= \frac{1}{\varepsilon} (v_i \mathbf{n} \cdot \mathbf{e}_i) \frac{\partial \varphi}{\partial \eta} + v_i \left( \frac{\partial \varphi}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right) \\ &= \frac{1}{\varepsilon} \mathbf{v} \cdot \mathbf{n} \frac{\partial \varphi}{\partial \eta} + v_i \left( \frac{\partial \varphi}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right). \end{aligned}$$

où on utilise le résultat standard de géométrie différentielle :  $\frac{\partial \eta}{\partial x_i} = \frac{1}{\varepsilon} \mathbf{n} \cdot \mathbf{e}_i$ .

En définissant les composantes normale et tangentielle de la vitesse microscopique

$$v_\eta = \mathbf{v} \cdot \mathbf{n} \quad , \quad v_\tau = \mathbf{v} \cdot \mathbf{t} \quad , \quad v_\chi = \mathbf{v} \cdot \mathbf{s},$$

on obtient :  $\mathbf{v} \cdot \nabla_x \varphi = \frac{1}{\varepsilon} v_\eta \frac{\partial \varphi}{\partial \eta} + v_i \left( \frac{\partial \varphi}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right)$ , et l'équation (2.33) se réécrit

$$\boxed{v_\eta \frac{\partial \varphi}{\partial \eta} = L(\varphi) - \varepsilon v_i \left( \frac{\partial \varphi}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right).}$$

En notant pour alléger l'écriture  $G|_\Gamma = G(\eta = 0, v_\eta > 0)$  pour toute grandeur  $G$ , on obtient la condition aux limites sur  $\varphi$  avec (2.28) :

$$f|_\Gamma = (\bar{f} + \varphi)|_\Gamma = p_w(f) M[1, \mathbf{0}, 1] \Leftrightarrow \varphi|_\Gamma = p_w(f) M[1, \mathbf{0}, 1] - \bar{f}|_\Gamma$$

(on ne développe pas  $p_w(f)$ , qui n'est pas linéaire). On obtient ainsi le problème de demi-espace

$$\boxed{\begin{cases} v_\eta \frac{\partial \varphi}{\partial \eta} = L(\varphi) - \varepsilon v_i \left( \frac{\partial \varphi}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right), & \eta > 0, \\ \varphi|_\Gamma = p_w(f) M[1, \mathbf{0}, 1] - \bar{f}|_\Gamma, & v_\eta > 0. \end{cases}} \quad (2.39)$$

L'introduction du développement (2.34) dans la première équation de (2.39) en tenant compte de (2.35) donne

$$\begin{aligned} \varepsilon v_\eta \frac{\partial \varphi_1}{\partial \eta} + \varepsilon^2 v_\eta \frac{\partial \varphi_2}{\partial \eta} + \dots &= \varepsilon L(\varphi_1) + \varepsilon^2 L(\varphi_2) + \dots \\ &\quad - \varepsilon^2 v_i \left( \frac{\partial \varphi_1}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi_1}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right) - \varepsilon^3 v_i \left( \frac{\partial \varphi_2}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi_2}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right) - \dots \end{aligned}$$

et par identification, on obtient

$$\begin{cases} v_\eta \frac{\partial \varphi_1}{\partial \eta} = L(\varphi_1), \\ v_\eta \frac{\partial \varphi_2}{\partial \eta} = L(\varphi_2) - v_i \left( \frac{\partial \varphi_1}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi_1}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right), \\ \vdots \\ v_\eta \frac{\partial \varphi_{n+1}}{\partial \eta} = L(\varphi_{n+1}) - v_i \left( \frac{\partial \varphi_n}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi_n}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right) \quad \forall n \in \mathbb{N}. \end{cases} \quad (2.40)$$

La condition aux limites donnée par (2.39) s'écrit

$$\varepsilon \varphi_1|_\Gamma + \varepsilon^2 \varphi_2|_\Gamma + \dots = \varepsilon p_w(f_1) M[1, \mathbf{0}, 1] + \varepsilon^2 p_w(f_2) M[1, \mathbf{0}, 1] + \dots - \varepsilon \bar{f}_1|_\Gamma - \varepsilon^2 \bar{f}_2|_\Gamma - \dots$$

soit

$$\varphi_n|_\Gamma = p_w(f_n) M[1, \mathbf{0}, 1] - \bar{f}_n|_\Gamma \quad \forall n \in \mathbb{N}^*.$$

Le correcteur n'agissant qu'en zone de proche paroi, on dispose également de la condition à l'infini :

$$\varphi(\eta, \tau, \chi, \mathbf{v}) \xrightarrow{\eta \rightarrow \infty} 0 \quad \text{soit} \quad \varphi_n(\eta, \tau, \chi, \mathbf{v}) \xrightarrow{\eta \rightarrow \infty} 0 \quad \forall n \in \mathbb{N}^*. \quad (2.41)$$



On dispose ainsi à chaque ordre d'un problème de demi-espace composé d'une équation, d'une condition de bord et d'une condition à l'infini. On va résoudre dans la partie suivante le problème vérifié par la composante  $\varphi_1$  qui va permettre d'aboutir aux conditions dites de "glissement".

☞ NB : on peut aussi obtenir un problème de demi-espace sur  $\varphi_0$  : si on ne tient pas compte de (2.35) dans le développement de Hilbert, on arrive à

$$\begin{cases} v_\eta \frac{\partial \varphi_0}{\partial \eta} = L(\varphi_0), \\ \varphi_0|_\Gamma = p_w(f_0)M[1, \mathbf{0}, 1] - \bar{f}_0|_\Gamma, \\ \varphi_0(\eta, \tau, \chi, \mathbf{v}) \xrightarrow{\eta \rightarrow \infty} 0. \end{cases}$$

Or on sait d'après (2.30) que  $\bar{f}_0 = \left[ \rho'_0 + \mathbf{u}_0' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_0 \right] M[1, \mathbf{0}, 1] = p_0(\mathbf{v})M[1, \mathbf{0}, 1]$ , donc :

$$\varphi_0|_\Gamma = \underbrace{\left[ p_w(f_0) - p_0(\mathbf{v}) \right]}_{= p_{\varphi_0}(\mathbf{v})} M[1, \mathbf{0}, 1]$$

On a alors la solution triviale  $\varphi_0 \equiv 0$  du système à condition que  $p_{\varphi_0}(\mathbf{v}) = 0$ , ce qui donne bien les conditions d'adhérence à la paroi (2.32).  $\square$

### 2.4.3.3 Correcteur au premier ordre : conditions de glissement

Le terme  $\varphi_1$  vérifie le système :

$$\begin{cases} v_\eta \frac{\partial \varphi_1}{\partial \eta} = L(\varphi_1), \\ \varphi_1|_\Gamma = p_w(f_1)M[1, \mathbf{0}, 1] - \bar{f}_1|_\Gamma, \\ \varphi_1(\eta, \tau, \chi, \mathbf{v}) \xrightarrow{\eta \rightarrow \infty} 0. \end{cases} \quad (2.42)$$

Or le système (2.30) donne

$$\bar{f}_1 = -\mathbf{v} \cdot \nabla_x \bar{f}_0 + \left[ \rho'_1 + \mathbf{u}_1' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_1 \right] M[1, \mathbf{0}, 1] = -\mathbf{v} \cdot \nabla_x \bar{f}_0 + p_1(\mathbf{v})M[1, \mathbf{0}, 1],$$

avec  $\bar{f}_0 = \left[ \rho'_0 + \mathbf{u}_0' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_0 \right] M[1, \mathbf{0}, 1] = p_0(\mathbf{v})M[1, \mathbf{0}, 1]$  selon (2.21).

Le terme  $\mathbf{v} \cdot \nabla_x \bar{f}_0$  fait apparaître un polynôme du 3<sup>e</sup> degré en vitesses dont le coefficient dominant dépend de  $f_0$ ; un calcul simple donne en effet

$$\begin{aligned}
-\mathbf{v} \cdot \nabla_x \bar{f}_0 &= -\mathbf{v} \cdot \nabla_x p_0(\mathbf{v}) M[1, \mathbf{0}, 1] = -\mathbf{v} \cdot \nabla_x \left[ \rho'_0 + \mathbf{u}_0' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_0 \right] M[1, \mathbf{0}, 1] \\
&= -\mathbf{v} \cdot \left[ \nabla_x \rho'_0 + (\nabla_x \mathbf{u}_0')^T \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \nabla_x T'_0 \right] M[1, \mathbf{0}, 1] \\
&= \left[ -\nabla_x \rho'_0 \cdot \mathbf{v} - \nabla \cdot [(\mathbf{v} \otimes \mathbf{v}) \mathbf{u}_0'] - \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \nabla_x T'_0 \cdot \mathbf{v} \right] M[1, \mathbf{0}, 1] \\
&= \left[ -\nabla \cdot [(\mathbf{v} \otimes \mathbf{v}) \mathbf{u}_0'] - \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \nabla_x T'_0 \cdot \mathbf{v} \right] M[1, \mathbf{0}, 1] \quad \text{selon (2.23);}
\end{aligned}$$

la condition aux limites du système (2.42) s'écrit alors de façon détaillée

$$\begin{aligned}
\varphi_1|_\Gamma &= \left[ p_w(f_1) - p_1(\mathbf{v}) + \mathbf{v} \cdot \nabla_x p_0(\mathbf{v}) \right] M[1, \mathbf{0}, 1] \\
&= \left[ \sigma_w(f_1) - \rho'_1 - \mathbf{u}_1' \cdot \mathbf{v} - T'_1 \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) + \nabla \cdot [(\mathbf{v} \otimes \mathbf{v}) \mathbf{u}_0'] + \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \nabla_x T'_0 \cdot \mathbf{v} \right] M[1, \mathbf{0}, 1];
\end{aligned}$$

il n'y a donc pas de solution triviale au problème. On va cependant pouvoir estimer les valeurs au bord des quantités fluides relatives à l'ordre 1, soit les trois moments de  $\bar{f}_1$  :  $\rho'_1$ ,  $\mathbf{u}_1'$  et  $T'_1$ .

La valeur de la vitesse normale s'obtient aisément. On a en intégrant la première équation de (2.42) par rapport à  $\mathbf{v}$  :

$$\left\langle v_\eta \frac{\partial \varphi_1}{\partial \eta} \right\rangle = \left\langle L(\varphi_1) \right\rangle = 0.$$

Or :

$$\left\langle v_\eta \frac{\partial \varphi_1}{\partial \eta} \right\rangle = \left\langle \mathbf{v} \cdot \mathbf{n} \frac{\partial \varphi_1}{\partial \eta} \right\rangle = \frac{\partial}{\partial \eta} \left( \langle \mathbf{v} \cdot \mathbf{n} \varphi_1 \rangle \right) = \frac{\partial}{\partial \eta} \left( \langle \mathbf{v} \varphi_1 \rangle \cdot \mathbf{n} \right);$$

la quantité  $\langle \mathbf{v} \varphi_1 \rangle \cdot \mathbf{n}$  est donc uniforme. Or d'après la condition à l'infini (2.41),  $\langle \mathbf{v} \varphi_1 \rangle \cdot \mathbf{n} \xrightarrow{\eta \rightarrow \infty} 0$ , et donc  $\langle \mathbf{v} \varphi_1 \rangle \cdot \mathbf{n} = 0$  en tout point du domaine de calcul. En notant  $\mathbf{u}_{\varphi_1}' = \langle \mathbf{v} \varphi_1 \rangle$  par analogie avec (2.17), on obtient :  $\mathbf{u}_{\varphi_1}' \cdot \mathbf{n} = 0$ . On a alors :

$$\begin{aligned}
\langle \mathbf{v} f_1 \rangle \cdot \mathbf{n} &= \langle \mathbf{v} \bar{f}_1 \rangle \cdot \mathbf{n} + \underbrace{\langle \mathbf{v} \varphi_1 \rangle \cdot \mathbf{n}}_{=0} = \underbrace{\langle \mathbf{v} (-\mathbf{v} \cdot \nabla_x \bar{f}_0) \rangle \cdot \mathbf{n}}_{= \langle \mathbf{v} L(\bar{f}_1) \rangle = \mathbf{0}} + \underbrace{\langle \mathbf{v} p_1(\mathbf{v}) M[1, \mathbf{0}, 1] \rangle \cdot \mathbf{n}}_{= \mathbf{u}_1'} = \mathbf{u}_1' \cdot \mathbf{n}.
\end{aligned}$$

Le système (2.31) donne :  $\langle \mathbf{v} f_1 \rangle_\Gamma = \mathbf{0}$ ; on a donc  $\mathbf{u}_1' \cdot \mathbf{n} = 0$  sur  $\Gamma$ .

On obtient ainsi les conditions aux limites sur la vitesse normale :

$$\boxed{
\begin{cases}
\mathbf{u}_1' \cdot \mathbf{n} = 0, \\
\mathbf{u}_{\varphi_1}' \cdot \mathbf{n} = 0.
\end{cases}
} \quad (2.43)$$

Les autres quantités s'obtiennent de façon plus complexe, en analysant la compatibilité entre la condition de bord et la condition à l'infini du système (2.42). En effet, supposons qu'il existe  $\varphi_\infty \in L^2(\mathbb{R}^+ \times \mathbb{R}^2 \times \mathbb{R}^3)$  tel que  $\varphi_1(\eta, \tau, \chi, \mathbf{v}) \xrightarrow{\eta \rightarrow \infty} \varphi_\infty(\eta, \tau, \chi, \mathbf{v})$ . On a alors  $v_\eta \frac{\partial \varphi_1}{\partial \eta} \xrightarrow{\eta \rightarrow \infty} 0$ , et selon l'équation (2.42), on a alors  $L(\varphi_\infty) = 0$ , et  $\varphi_\infty$  est donc une maxwellienne linéarisée. De plus,  $\langle v_\eta \varphi_\infty \rangle = 0$ , donc  $\varphi_\infty$  s'écrit sous la forme

$$\varphi_\infty = \left[ a + b_\tau v_\tau + b_\chi v_\chi + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) c \right] M[1, \mathbf{0}, 1], \quad a, b_\tau, b_\chi, c \in \mathbb{R}. \quad (2.44)$$

(Pour une démonstration plus rigoureuse de ce résultat, voir l'étude effectuée par Bardos, Caflisch et Nikolaenko dans [11]). En calculant la limite asymptotique du problème, puis en l'identifiant à  $\varphi_\infty$ , on obtient des conditions sur les coefficients de  $\varphi_1|_\Gamma$  pour obtenir  $\varphi_\infty = 0$ , c'est-à-dire  $a = b_\tau = b_\chi = c = 0$ .

La limite asymptotique  $\varphi_\infty$  dépendant linéairement de  $\varphi_1|_\Gamma$ , on peut décomposer le problème initial en sous-problèmes simples en considérant séparément les termes de  $\varphi_1|_\Gamma$ . On réécrit ainsi la valeur de  $\varphi_1|_\Gamma$ , dans laquelle on va distinguer trois parties :

$$\underbrace{\left[ \sigma_w(f_1) - \rho'_1 - \mathbf{u}_1' \cdot \mathbf{v} - T'_1 \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \right]}_{(1)} + \underbrace{\nabla \cdot [(\mathbf{v} \otimes \mathbf{v}) \mathbf{u}_0']}_{(2)} + \underbrace{\left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \nabla_x T'_0 \cdot \mathbf{v}}_{(3)} M[1, \mathbf{0}, 1].$$

où selon l'équation (2.43)

$$\mathbf{u}_1' \cdot \mathbf{v} = u'_{1\tau} v_\tau + u'_{1\chi} v_\chi,$$

et on calcule alors la limite asymptotique des termes (1), (2) et (3).

- D'après l'étude de Coron [29], la partie (1) est préservée dans la limite asymptotique.
- La partie (2) se réécrit de façon développée

$$\begin{aligned} \nabla \cdot [(\mathbf{v} \otimes \mathbf{v}) \mathbf{u}_0'] &= \mathbf{v} \cdot \nabla (\mathbf{u}_0' \cdot \mathbf{v}) = \mathbf{v} \cdot \nabla (u'_{0\eta} v_\eta + u'_{0\tau} v_\tau + u'_{0\chi} v_\chi) \\ &= v_\eta \mathbf{v} \cdot \nabla (u'_{0\eta}) + v_\tau \mathbf{v} \cdot \nabla (u'_{0\tau}) + v_\chi \mathbf{v} \cdot \nabla (u'_{0\chi}) \\ &= v_\eta \left( v_\eta \frac{\partial u'_{0\eta}}{\partial n} + v_\tau \frac{\partial u'_{0\eta}}{\partial \tau} + v_\chi \frac{\partial u'_{0\eta}}{\partial \chi} \right) + v_\tau \left( v_\eta \frac{\partial u'_{0\tau}}{\partial n} + v_\tau \frac{\partial u'_{0\tau}}{\partial \tau} + v_\chi \frac{\partial u'_{0\tau}}{\partial \chi} \right) \\ &\quad + v_\chi \left( v_\eta \frac{\partial u'_{0\chi}}{\partial n} + v_\tau \frac{\partial u'_{0\chi}}{\partial \tau} + v_\chi \frac{\partial u'_{0\chi}}{\partial \chi} \right) \\ &= v_\tau v_\eta \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) + v_\chi v_\eta \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right) + v_\tau v_\chi \left( \frac{\partial u'_{0\tau}}{\partial \chi} + \frac{\partial u'_{0\chi}}{\partial \tau} \right) \\ &\quad + v_\eta^2 \frac{\partial u'_{0\eta}}{\partial n} + v_\tau^2 \frac{\partial u'_{0\tau}}{\partial \tau} + v_\chi^2 \frac{\partial u'_{0\chi}}{\partial \chi}, \end{aligned}$$

en notant  $\frac{\partial}{\partial n}$  la dérivée normale et  $\frac{\partial}{\partial \tau}$  et  $\frac{\partial}{\partial \chi}$  les dérivées tangentiellles dans les directions  $\mathbf{t}$  et  $\mathbf{s}$  respectivement.

☞ NB : d'après la définition de  $\eta$  (2.38),  $\frac{\partial}{\partial \eta} = \varepsilon \frac{\partial}{\partial n}$ .  $\square$

Les 4 derniers termes de l'expression obtenue ci-dessus sont nuls. En effet, selon la condition d'adhérence (2.32), les vecteurs  $\mathbf{u}_{\mathbf{0}'}'$  et  $\mathbf{u}_{\mathbf{w}}$  étant égaux de part et d'autre de la paroi, leurs dérivées suivant les directions tangentes à cette paroi sont égales :

$$\frac{\partial u'_{0*}}{\partial \tau} = \frac{\partial u_{w*}}{\partial \tau} \text{ et } \frac{\partial u'_{0*}}{\partial \chi} = \frac{\partial u_{w*}}{\partial \chi}, \quad * \in \{\eta, \tau, \chi\}.$$

La paroi étant supposée indéformable,

$$\frac{\partial u_{w\tau}}{\partial \tau} = \frac{\partial u_{w\chi}}{\partial \chi} = \left( \frac{\partial u_{w\tau}}{\partial \chi} + \frac{\partial u_{w\chi}}{\partial \tau} \right) = 0,$$

ce qui implique

$$\frac{\partial u'_{0\tau}}{\partial \tau} = \frac{\partial u'_{0\chi}}{\partial \chi} = \left( \frac{\partial u'_{0\tau}}{\partial \chi} + \frac{\partial u'_{0\chi}}{\partial \tau} \right) = 0.$$

Or  $\nabla \cdot \mathbf{u}_{\mathbf{0}'}' = \frac{\partial u'_{0\eta}}{\partial n} + \frac{\partial u'_{0\tau}}{\partial \tau} + \frac{\partial u'_{0\chi}}{\partial \chi} = 0$  selon (2.22), et on a donc  $\frac{\partial u'_{0\eta}}{\partial n} = 0$ .

En revanche, les dérivées normales étant calculées "dans le fluide" et non "dans le solide", les termes

$$\left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) \text{ et } \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right)$$

sont a priori non nuls. Il reste donc

$$\nabla \cdot [(\mathbf{v} \otimes \mathbf{v}) \mathbf{u}_{\mathbf{0}'}'] = v_{\tau} v_{\eta} \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) + v_{\chi} v_{\eta} \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right).$$

On calcule la limite de ce terme en exploitant les propriétés de symétrie de l'équation. En effet, on constate que celle-ci préserve la symétrie par rapport à  $v_{\tau}$  et  $v_{\chi}$  (i.e. si  $\varphi_1$  est solution, les fonctions  $\zeta_1(\eta, v_{\eta}, v_{\tau}, v_{\chi}) = \varphi_1(\eta, v_{\eta}, -v_{\tau}, v_{\chi})$  et  $\zeta_2(\eta, v_{\eta}, v_{\tau}, v_{\chi}) = \varphi_1(\eta, v_{\eta}, v_{\tau}, -v_{\chi})$  sont également solutions). Ceci implique que les propriétés de parité ou d'imparité des différents termes de la condition de bord par rapport à ces deux variables sont préservées dans leur limite asymptotique : en effet, si  $\varphi_1$  est la solution correspondant à une donnée au bord paire par rapport à  $v_{\tau}$ , alors la préservation de la symétrie implique  $\zeta_1$  solution, et la parité de la condition au bord indique que  $\zeta_1$  a la même donnée au bord que  $\varphi$  ; par unicité,  $\zeta_1 \equiv \varphi$ , et  $\varphi$  est donc paire par rapport à  $v_{\tau}$  (un raisonnement analogue donne la même chose pour  $v_{\chi}$  et pour une condition de bord impaire). Cette limite étant nécessairement de la même forme que  $\varphi_{\infty}$ , on peut déduire son expression en analysant ces propriétés :

➤ le terme  $v_{\tau} v_{\eta} \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right)$  est impair par rapport à  $v_{\tau}$  et pair par rapport à  $v_{\chi}$ , on en déduit :

$$\exists C_1 \in \mathbb{R} / v_{\tau} v_{\eta} \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) M[1, \mathbf{0}, 1] \xrightarrow{\eta \rightarrow \infty} C_1 v_{\tau} \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) M[1, \mathbf{0}, 1] ; \quad (2.45)$$

- les variables  $v_\tau$  et  $v_\chi$  ont un rôle symétrique dans l'équation : leur permutation dans  $\varphi_{1|\Gamma}$  se répercute dans  $\varphi_\infty$ . On a donc

$$v_\chi v_\eta \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right) M[1, \mathbf{0}, 1] \xrightarrow{\eta \rightarrow \infty} C_1 v_\chi \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right) M[1, \mathbf{0}, 1],$$

soit

$$\nabla \cdot \left[ (\mathbf{v} \otimes \mathbf{v}) \mathbf{u}_0' \right] \xrightarrow{\eta \rightarrow \infty} C_1 \left[ v_\tau \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) + v_\chi \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right) \right] M[1, \mathbf{0}, 1].$$

☞ NB : même sans montrer que le terme  $v_\tau v_\chi \left( \frac{\partial u'_{0\tau}}{\partial \chi} + \frac{\partial u'_{0\chi}}{\partial \tau} \right)$  est nul, on voit facilement qu'il n'influe pas sur la limite asymptotique. En effet, ce terme est impair par rapport à  $v_\tau$  et à  $v_\chi$ , or l'expression de  $\varphi_\infty$  ne comprend aucun terme avec cette double imparité .  $\square$

- On réécrit également la partie (3) :

$$\begin{aligned} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \nabla_x T_0' \cdot \mathbf{v} &= \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \left( v_\eta \frac{\partial T_0'}{\partial n} + v_\tau \frac{\partial T_0'}{\partial \tau} + v_\chi \frac{\partial T_0'}{\partial \chi} \right) \\ &= v_\eta \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \frac{\partial T_0'}{\partial n} + v_\tau \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \frac{\partial T_0'}{\partial \tau} + v_\chi \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \frac{\partial T_0'}{\partial \chi}, \end{aligned}$$

et on a de la même façon

$$\exists C_2, C_3 \in \mathbb{R} / v_\eta \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \frac{\partial T_0'}{\partial n} M[1, \mathbf{0}, 1] \xrightarrow{\eta \rightarrow \infty} \left[ C_2 + C_3 \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \right] \frac{\partial T_0'}{\partial n} M[1, \mathbf{0}, 1], \quad (2.46)$$

$$\exists C_4 \in \mathbb{R} / \begin{cases} v_\tau \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \frac{\partial T_0'}{\partial \tau} M[1, \mathbf{0}, 1] \xrightarrow{\eta \rightarrow \infty} C_4 v_\tau \frac{\partial T_0'}{\partial \tau} M[1, \mathbf{0}, 1], \\ v_\chi \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) \frac{\partial T_0'}{\partial \chi} M[1, \mathbf{0}, 1] \xrightarrow{\eta \rightarrow \infty} C_4 v_\chi \frac{\partial T_0'}{\partial \chi} M[1, \mathbf{0}, 1]. \end{cases} \quad (2.47)$$

Les limites (2.45), (2.46) et (2.47) donnent finalement par identification avec (2.44) :

$$\begin{cases} a = \sigma_w(f_1) - \rho_1' + C_2 \frac{\partial T_0'}{\partial n}, \\ b_\tau = -u_{1\tau}' + C_1 \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) + C_4 \frac{\partial T_0'}{\partial \tau}, \\ b_\chi = -u_{1\chi}' + C_1 \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right) + C_4 \frac{\partial T_0'}{\partial \chi}, \\ c = -T_1' + C_3 \frac{\partial T_0'}{\partial n}. \end{cases}$$

La condition à l'infini de (2.42) impose  $a = b_\tau = b_\chi = c = 0$ , soit en renommant les constantes

$$\boxed{\begin{cases} \rho'_1 - \sigma_w(f_1) = K_1 \frac{\partial T'_0}{\partial n}, \\ u'_{1\tau} = K_2 \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) + K_3 \frac{\partial T'_0}{\partial \tau}, \\ u'_{1\chi} = K_2 \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right) + K_3 \frac{\partial T'_0}{\partial \chi}, \\ T'_1 = K_4 \frac{\partial T'_0}{\partial n}. \end{cases}} \quad (2.48)$$

Les constantes  $K_1, K_2, K_3, K_4$ , appelées coefficients de glissement (ou de saut), sont les solutions des 3 problèmes de demi-espace

$$\begin{cases} v_\eta \frac{\partial \psi}{\partial \eta} = L(\psi), \\ \psi|_\Gamma = v_\tau v_\eta M[1, \mathbf{0}, 1], \\ \psi \xrightarrow{\eta \rightarrow \infty} K_2 v_\tau M[1, \mathbf{0}, 1], \end{cases}, \quad \begin{cases} v_\eta \frac{\partial \psi}{\partial \eta} = L(\psi), \\ \psi|_\Gamma = v_\tau \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1], \\ \psi \xrightarrow{\eta \rightarrow \infty} K_3 v_\tau M[1, \mathbf{0}, 1], \end{cases}$$

$$\begin{cases} v_\eta \frac{\partial \psi}{\partial \eta} = L(\psi), \\ \psi|_\Gamma = v_\eta \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1], \\ \psi \xrightarrow{\eta \rightarrow \infty} \left[ K_1 + K_4 \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \right] M[1, \mathbf{0}, 1]. \end{cases}$$

On les calcule numériquement, avec un modèle réduit à 2 vitesses microscopiques (décrit dans l'annexe C). Les valeurs obtenues sont :

$$\begin{cases} K_1 = 0.9972, \\ K_2 = 1.4390, \\ K_3 = 0.7477, \\ K_4 = 1.8738. \end{cases}$$

Ces valeurs sont sensiblement différentes de celles données par Sone dans [63] et [64], qui sont

$$\begin{cases} K_1 = 0.3477, \\ K_2 = 1.0162, \\ K_3 = 0.3832, \\ K_4 = 1.3027. \end{cases}$$

La méthode utilisée par Sone pour calculer la valeur des constantes n'étant pas explicitée, nous ne pouvons pas donner d'explication à cette différence. On peut cependant supposer qu'elle provient d'une différence dans l'adimensionnement du problème. On discute dans la section suivante le signe de la constante  $K_3$  sur un cas simple.

Les systèmes (2.32) et (2.48) permettent de corriger les conditions aux limites du régime fluide lorsque l'effet du nombre de Knudsen n'est plus négligeable (régime de glissement). Pour la température, on a ainsi

$$T'_{|\Gamma} = T'_{0|\Gamma} + \varepsilon T'_{1|\Gamma} = \theta_w + \varepsilon K_4 \frac{\partial T'_0}{\partial n}_{|\Gamma} = \theta_w + \varepsilon K_4 \left( \frac{\partial T'}{\partial n}_{|\Gamma} - \varepsilon \frac{\partial T'_1}{\partial n}_{|\Gamma} \right) = \theta_w + \varepsilon K_4 \frac{\partial T'}{\partial n}_{|\Gamma} + \mathcal{O}(\varepsilon^2),$$

soit

$$\boxed{T'_{|\Gamma} - \varepsilon K_4 \frac{\partial T'}{\partial n}_{|\Gamma} = \theta_w,}$$

et de même pour la vitesse :

$$\boxed{\begin{cases} u'_{\tau|\Gamma} - \varepsilon K_2 \left( \frac{\partial u'_{\tau}}{\partial n} + \frac{\partial u'_{\eta}}{\partial \tau} \right) - \varepsilon K_3 \frac{\partial T'}{\partial \tau} = \mathbf{u}_w \cdot \mathbf{t}, \\ u'_{\chi|\Gamma} - \varepsilon K_2 \left( \frac{\partial u'_{\chi}}{\partial n} + \frac{\partial u'_{\eta}}{\partial \chi} \right) - \varepsilon K_3 \frac{\partial T'}{\partial \chi} = \mathbf{u}_w \cdot \mathbf{s}. \end{cases}}$$

Les conditions de glissement obtenues sont de type Robin avec une dérivée d'ordre 1 en  $\varepsilon$ .

On voit donc apparaître un "saut" sur les grandeurs macroscopiques : la vitesse et la température que l'on aurait au niveau de la paroi sans le correcteur sont différentes de la vitesse et de la température de paroi. Le principe de cette "condition de glissement" est illustré sur la figure (2.3) : pour une grandeur  $G$ , on représente en (1) la solution de Grad-Hilbert, dont la valeur à la paroi est notée  $\bar{G}_{|\Gamma}$ , et en (2) la solution réelle ; la différence  $G_s = G_{|\Gamma} - \bar{G}_{|\Gamma}$  est appelée la grandeur "de glissement". La zone de divergence des deux solutions, en (3), est la couche de Knudsen, conformément aux résultats obtenus au paragraphe 2.4.3.1.

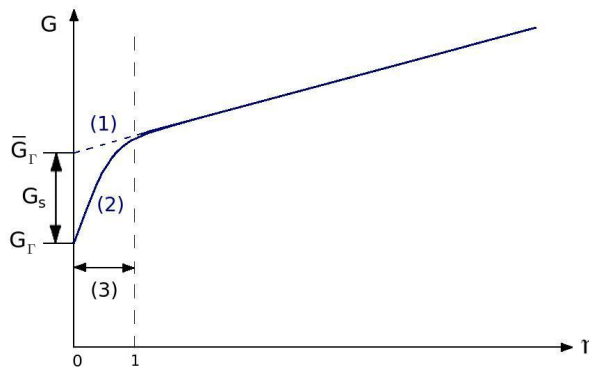


FIGURE 2.3: Condition de saut sur la grandeur  $G$

☞ NB : le coefficient  $K_1$  n'est pas réellement significatif : il n'y a pas de réelle condition aux limites sur la densité, la densité à la paroi étant déterminée par une condition de flux de masse nul à travers cette paroi. On verra au chapitre suivant les conséquences sur les problèmes limites déterminant la densité  $\rho'_0$  et la vitesse  $\mathbf{u}'_0$ . □

Pour  $n \geq 2$ , les problèmes de demi-espace sont plus complexes à traiter à cause des termes inhomogènes  $v_i \left( \frac{\partial \varphi_{n-1}}{\partial \tau} \frac{\partial \tau}{\partial x_i} + \frac{\partial \varphi_{n-1}}{\partial \chi} \frac{\partial \chi}{\partial x_i} \right)$ . Nous n'irons pas plus loin que le premier ordre.

#### 2.4.3.4 Glissement thermique : discussion du signe de la constante $K_3$

Le glissement thermique est un phénomène caractéristique de la dynamique des gaz raréfiés. Également appelé "transpiration thermique", il a été étudié par Reynolds [61], Maxwell [50] et Knudsen [44].

Son principe est le suivant : on considère un canal dans lequel circule un gaz raréfié. Les parois du canal sont soumises à un gradient de température uniforme, comme indiqué sur la figure 2.4. Ce gradient de température se retrouve par conduction dans le gaz. Dans la partie droite du canal, le gaz chaud possède une agitation thermique élevée ; dans la partie gauche, le gaz plus froid est moins agité. La résultante des chocs des particules de gaz sur la paroi est donc plus importante à droite qu'à gauche ; en conséquence, le gaz exerce une force sur la paroi, dirigée de la droite vers la gauche. La paroi étant immobile, la réaction exercée par la paroi sur le gaz met celui-ci en mouvement de la gauche vers la droite, c'est-à-dire de la zone froide vers la zone chaude, ce qui s'oppose *a priori* aux lois de la thermodynamique.

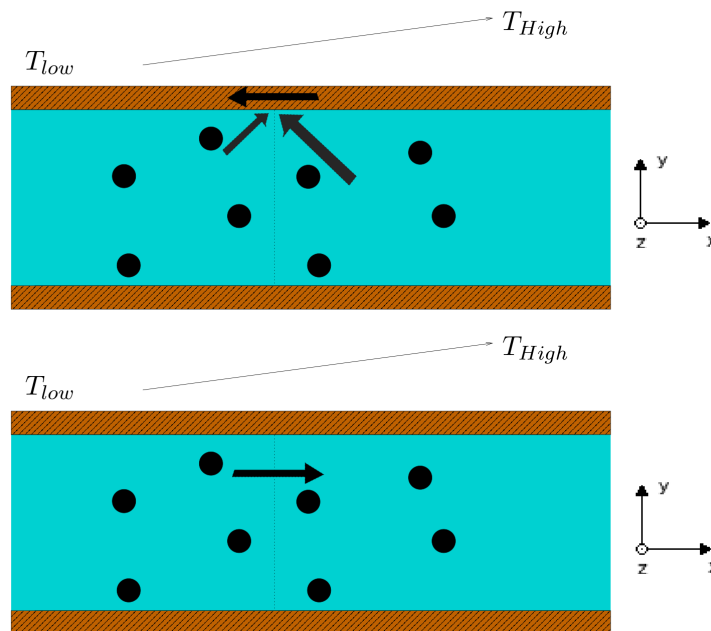


FIGURE 2.4: Glissement thermique : force exercée par le gaz sur la paroi (haut) et réaction (bas)



On analyse sur cet exemple les conditions de glissement obtenues au paragraphe précédent. On suppose  $\mathbf{u}_0' = \mathbf{0}$ , et on remarque que le gradient de température à la paroi est tangentiel, ce qui se traduit avec la condition d'adhérence (2.32) :  $\frac{\partial T_0'}{\partial n} = 0$ . Les relations (2.48) s'écrivent alors, en assimilant localement la direction tangentielle  $\tau$  à  $x$  (c'est-à-dire en posant  $\mathbf{t} = \mathbf{x}$  en chaque point de la paroi) :

$$\begin{cases} \rho_1' - \sigma_w(f_1) = 0, \\ u_{1\tau}' = K_3 \frac{\partial T_0'}{\partial x}, \\ T_1' = 0. \end{cases}$$

Le gaz se déplaçant de la gauche vers la droite, on a  $u_{1\tau}' \geq 0$ . De plus, avec la condition d'adhérence, on a  $\frac{\partial T_0'}{\partial x} = \frac{\partial T_w}{\partial x} \geq 0$ . Cela donne donc

$$K_3 \geq 0.$$

## 2.5 Bilan

Après avoir adimensionné l'équation BGK stationnaire, ce qui a permis de définir le nombre de Knudsen  $\text{Kn}$ , paramètre pertinent dans l'étude des gaz raréfiés, on a considéré un état proche de l'équilibre pour des vitesses très basses :

$$F = M[1, \mathbf{0}, 1] + f, \quad f \ll 1,$$

ce qui a permis d'effectuer une linéarisation de l'équation :

$$\mathbf{v} \cdot \nabla_x f = \frac{1}{\text{Kn}} L(f), \quad L(f) = \left[ \langle f \rangle + \langle \mathbf{v} f \rangle \cdot \mathbf{v} + \left\langle \frac{|\mathbf{v}|^2 - 3}{3} f \right\rangle \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \right] M[1, \mathbf{0}, 1] - f.$$

On a ensuite effectué un développement de Grad-Hilbert sur la perturbation  $f$ , en se plaçant à bas nombre de Knudsen ( $\text{Kn} = \varepsilon$ ) :

$$f(\varepsilon) = f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \dots$$

ce qui, une fois réinjecté dans l'équation linéarisée, nous a donné les équations de Stokes :

$$\nabla (\rho_0' + T_0') = \mathbf{0},$$

$$\begin{cases} \nabla \cdot \mathbf{u}_n' = 0, \\ \nabla (\rho_{n+1}' + T_{n+1}') = \Delta (\mathbf{u}_n'), \quad \forall n \in \mathbb{N}. \\ \Delta (T_n') = 0. \end{cases}$$

On a ensuite adjoint des conditions aux limites à ce système en considérant une réflexion purement diffuse aux parois; la température  $T_w = 1 + \theta_w$  et la vitesse  $\mathbf{u}_w$  des parois sont supposées proches de celles de l'équilibre du gaz afin que les perturbations engendrées dans le fluide soient infinitésimales.

La linéarisation de la condition de réflexion diffuse a alors montré qu'il était nécessaire d'ajouter un correcteur  $\varphi$  au développement de Grad-Hilbert pour satisfaire cette condition :

$$f(\varepsilon) = f_0 + \varphi_0 + \varepsilon(f_1 + \varphi_1) + \varepsilon^2(f_2 + \varphi_2) + \dots$$

Avec cette nouvelle décomposition, on a pu écrire les problèmes de demi-espace satisfaits par chacune des composantes de  $\varphi$ , qui nous ont permis d'aboutir aux conditions d'adhérence (ordre 0)

$$\begin{cases} \mathbf{u}_0|_{\Gamma'} = \mathbf{u}_w, \\ T_0'|_{\Gamma} = \theta_w, \end{cases}$$

et aux conditions de glissement (ordre 1)

$$\begin{cases} u'_{1\eta}|_{\Gamma} = 0 \\ u'_{1\tau}|_{\Gamma} = K_2 \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) + K_3 \frac{\partial T'_0}{\partial \tau}, \\ u'_{1\chi}|_{\Gamma} = K_2 \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right) + K_3 \frac{\partial T'_0}{\partial \chi}, \\ T'_{1|\Gamma} = K_4 \frac{\partial T'_0}{\partial n}, \end{cases}$$

qui permettent de corriger les conditions aux limites du régime fluide lorsque l'effet du nombre de Knudsen n'est plus négligeable (régime de glissement) :

$$\begin{cases} T'_{|\Gamma} - \theta_w = \varepsilon K_4 \frac{\partial T'}{\partial n}|_{\Gamma}, \\ u'_{\eta}|_{\Gamma} = \mathbf{u}_w \cdot \mathbf{n} = 0, \\ (\mathbf{u}'_{|\Gamma} - \mathbf{u}_w) \cdot \mathbf{t} = \varepsilon \left[ K_2 \left( \frac{\partial u'_{\tau}}{\partial n} + \frac{\partial u'_{\eta}}{\partial \tau} \right) + K_3 \frac{\partial T'}{\partial \tau} \right], \\ (\mathbf{u}'_{|\Gamma} - \mathbf{u}_w) \cdot \mathbf{s} = \varepsilon \left[ K_2 \left( \frac{\partial u'_{\chi}}{\partial n} + \frac{\partial u'_{\eta}}{\partial \chi} \right) + K_3 \frac{\partial T'}{\partial \chi} \right]. \end{cases}$$

avec

$$\begin{cases} K_2 = 1.4390, \\ K_3 = 0.7477, \\ K_4 = 1.8738. \end{cases}$$

On souhaite désormais construire des schémas numériques préservant l'ensemble des propriétés asymptotiques du problème BGK linéarisé décrites dans ce chapitre, c'est-à-dire le jeu d'équations de Stokes aux deux premiers ordres, avec les conditions aux limites correspondantes d'adhérence et de glissement. Dans le chapitre suivant, on étudie la préservation de ces propriétés par les schémas Galerkin Discontinus.



## Chapitre 3

# Application de la méthode Galerkin Discontinu au problème BGK linéarisé avec condition aux limites de réflexion diffuse

*Le mystérieux est rarement discret, le curieux ne l'est jamais.*

Sosthène de La Rochefoucauld- *Pensées*

De nombreux travaux ont été effectués ces dernières années sur les schémas numériques dits "Asymptotic-Preserving" (AP). Ces schémas, consistants avec la limite asymptotique de l'équation cinétique discrétisée, possèdent la particularité de pouvoir s'affranchir de la contrainte de résolution de la couche limite cinétique, ce qui permet de ne pas raffiner excessivement le maillage au bord du domaine de calcul.

Les schémas Galerkin Discontinu, introduits au début des années 1970 par Reed et Hill pour la résolution numérique de l'équation de transport des neutrons [60] et qui connaissent un développement exponentiel depuis la fin des années 1990, ont montré de bonnes propriétés dans ce domaine. Pour le transport des neutrons en particulier, les travaux de Larsen, Morel et Miller [48], puis ceux de Guermond et Kanschat [37] ont montré que les schémas Galerkin Discontinu sont "AP" jusqu'à la couche limite forte, et sont ainsi capables de capturer correctement la solution asymptotique du problème limite de diffusion avec condition de Dirichlet sur le bord. Dans ce chapitre, nous étudions l'application de ces schémas au problème BGK linéarisé étudié dans le chapitre précédent, afin de voir s'ils possèdent des propriétés semblables pour la dynamique des gaz (dans laquelle la couche limite est faible).

Les schémas Galerkin Discontinu peuvent être considérés de deux façons. On peut les voir comme des schémas Éléments Finis autorisant une discontinuité sur les interfaces du maillage, c'est-à-dire comme des éléments finis "locaux". Chaque maille étant alors reliée aux autres par des flux numériques aux interfaces, on peut aussi considérer la méthode Galerkin Discontinu comme une extension de la méthode des Volumes Finis, dans laquelle la solution est approchée sur chaque maille par une fonction polynomiale de degré quelconque, et non plus simplement par une constante. L'analyse mathématique des schémas Galerkin Discontinu effectuée dans ce chapitre se fonde sur la première optique, bien adaptée à cette étude (on verra au chapitre 4 que

l'implémentation de ces schémas est très commode en les considérant plutôt comme des schémas de type Volumes Finis).

La démarche suivie est similaire à celle de Guermond et Kanschat. Il y a cependant des différences notables entre l'équation de transport linéaire étudiée par ces deux auteurs et le problème BGK linéarisé qui nous intéresse. La première est la dimension de l'espace des invariants : celui-ci est de dimension 1 dans le cas du transport linéaire, de dimension 5 pour l'équation BGK. En conséquence, l'équation de transport linéaire donne une seule équation limite de diffusion portant sur la même variable que l'équation cinétique, alors que l'équation BGK permet d'aboutir à deux équations scalaires (une équation de continuité et une équation de Laplace) et une équation vectorielle (équation de Stokes) liant les 5 premiers moments de la fonction de distribution. La seconde différence essentielle se trouve dans la nature du noyau de l'opérateur de collision, composé dans les deux cas de fonctions aux propriétés mathématiques très différentes : dans le cas du transport linéaire, le noyau se compose des fonctions isotropes, c'est-à-dire indépendantes de la variable angulaire (analogue de la variable de vitesse dans le cas de BGK), alors que le noyau de l'opérateur BGK se compose des fonctions maxwelliennes linéarisées. Certaines propriétés très simples dans le cas du transport linéaire deviennent ainsi beaucoup plus complexes à démontrer pour l'équation BGK (on le verra en particulier lors de la preuve de la continuité de la première composante du développement de Hilbert de la fonction de distribution discrète à la section 3.4.3.1).

Après avoir rappelé les problèmes limites de BGK linéarisé, nous étudions les conditions sous lesquelles ils sont bien posés. Nous détaillons ensuite la construction d'un schéma Galerkin Discontinu avec un flux numérique upwind, que nous appliquons au problème BGK linéarisé avec réflexion diffuse. Par un développement de Hilbert puis un choix adapté des fonctions-test, nous dérivons ensuite les schémas limites à l'ordre 0. Enfin, nous étudions ces problèmes et la convergence de leurs solutions.

Notons qu'un problème similaire est traité par Klar, de façon très différente, pour un schéma Différences Finies AP pour l'équation BGK linéaire dans [42].

### 3.1 Rappel du problème aux limites linéarisé

Le problème BGK linéarisé s'écrit :

$$\begin{cases} \mathbf{v} \cdot \nabla_x f(\mathbf{x}, \mathbf{v}) = \frac{1}{\varepsilon} L(f)(\mathbf{x}, \mathbf{v}), & \mathbf{v} \in \mathbb{R}^3, \mathbf{x} \in \Omega \subset \mathbb{R}^3 \\ f|_{\Gamma, \mathbf{v} \cdot \mathbf{n} > 0} = p_w(f) M[1, \mathbf{0}, 1](\mathbf{v}), & \Gamma = \partial\Omega \end{cases} \quad (3.1)$$

avec  $\varepsilon = \text{Kn}$ , et :

$$\begin{aligned} L(f) &= \left[ \langle f \rangle + \langle \mathbf{v} f \rangle \cdot \mathbf{v} + \left\langle \frac{|\mathbf{v}|^2 - 3}{3} f \right\rangle \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \right] M[1, \mathbf{0}, 1] - f \\ &= \left[ \rho'(\mathbf{x}) + \mathbf{u}'(\mathbf{x}) \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'(\mathbf{x}) \right] M[1, \mathbf{0}, 1](\mathbf{v}) - f(\mathbf{x}, \mathbf{v}) \\ &= \Pi(\rho', \mathbf{u}', T') - f, \end{aligned} \quad (3.2)$$

où

$$\Pi(\rho', \mathbf{u}', T')(\mathbf{x}, \mathbf{v}) = \left[ \rho'(\mathbf{x}) + \mathbf{u}'(\mathbf{x}) \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'(\mathbf{x}) \right] M[1, \mathbf{0}, 1](\mathbf{v})$$

et

$$\begin{cases} p_w(f) = \sigma_w(f) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w \\ \sigma_w(f) = -\frac{1}{2} \theta_w - \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f \right\rangle, \quad \mathbf{x} \in \Gamma \end{cases}$$

Le développement de Hilbert appliqué à l'équation et à la condition aux limites, avec l'introduction d'un correcteur de couche limite, ont permis d'établir l'expression des composantes de  $f$  à chaque ordre

$$\begin{cases} f_0 = \left[ \rho'_0 + \mathbf{u}_0' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_0 \right] M[1, \mathbf{0}, 1], \\ f_{n+1} = -(\mathbf{v} \cdot \nabla_x f_n) + \left[ \rho'_{n+1} + \mathbf{u}_{\mathbf{n}+1}' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T'_{n+1} \right] M[1, \mathbf{0}, 1] \quad \forall n \in \mathbb{N}, \end{cases}$$

ainsi que le jeu d'équations de Stokes :

$$\begin{aligned} \nabla(\rho'_0 + T'_0) &= \mathbf{0}, \\ \forall n \in \mathbb{N}, \quad &\begin{cases} \nabla \cdot \mathbf{u}_n' = 0, \\ \nabla(\rho'_{n+1} + T'_{n+1}) = \Delta(\mathbf{u}_n'), \\ \Delta(T'_n) = 0, \end{cases} \end{aligned}$$

les conditions aux limites d'adhérence pour les quantités relatives à l'ordre 0 du développement

$$\begin{cases} \rho'_0 = \sigma_w(f_0), \\ \mathbf{u}_0|_{\Gamma}' = \mathbf{u}_w, \\ T'_0|_{\Gamma} = \theta_w, \end{cases} \quad (3.3)$$

et enfin les conditions de glissement pour l'ordre suivant

$$\begin{cases} u'_{1\eta}|_{\Gamma} = 0, \\ u'_{1\tau}|_{\Gamma} = K_2 \left( \frac{\partial u'_{0\tau}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \tau} \right) + K_3 \frac{\partial T'_0}{\partial \tau}, \\ u'_{1\chi}|_{\Gamma} = K_2 \left( \frac{\partial u'_{0\chi}}{\partial n} + \frac{\partial u'_{0\eta}}{\partial \chi} \right) + K_3 \frac{\partial T'_0}{\partial \chi}, \\ T'_{1\Gamma} = K_4 \frac{\partial T'_0}{\partial n}. \end{cases} \quad \text{avec} \quad \begin{cases} K_2 = 1.4390, \\ K_3 = 0.7477, \\ K_4 = 1.8738. \end{cases} \quad (3.4)$$

Pour simplifier les écritures, on note dans tout ce chapitre  $M[1, \mathbf{0}, 1] = M$ .

### 3.2 Formulation variationnelle du jeu d'équations de Stokes avec conditions aux limites

Les équations et les conditions aux limites obtenues avec les deux premiers ordres du développement de Hilbert sur le problème continu permettent de déterminer les problèmes aux limites vérifiés par les champs  $\rho'_0$ ,  $\mathbf{u}_0'$ ,  $T'_0$ , comme indiqué sur la figure 3.1.

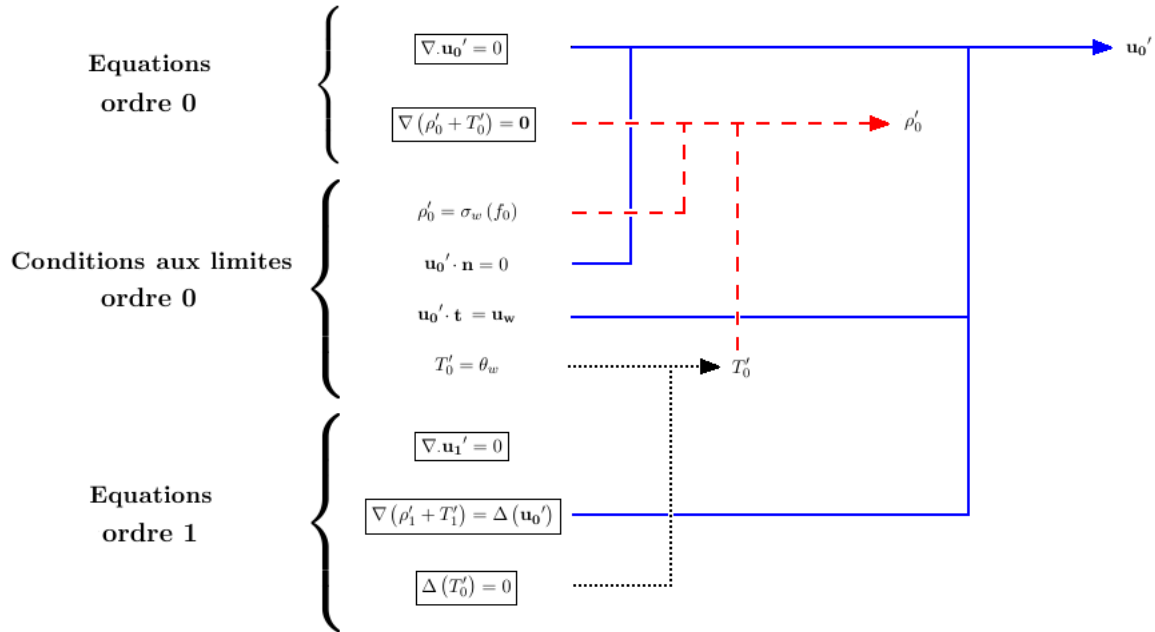


FIGURE 3.1: Équations macroscopiques et conditions aux limites

On montre ici que ces trois problèmes aux limites, que l'on souhaite préserver dans la limite asymptotique du problème BGK discret, sont bien posés. On utilise pour cela, de façon très classique, une formulation variationnelle de ces problèmes. Tous les outils d'analyse fonctionnelle utilisés dans ce paragraphe sont exposés dans la référence [17].

#### 3.2.1 Problème de Laplace : champ $T'_0$

Le problème de type Laplace non homogène donnant la température  $T'_0$  s'écrit

$$\begin{cases} \Delta T'_0 = 0, \\ T'_{0|\Gamma} = \theta_w. \end{cases}$$

La condition aux limites étant de type Dirichlet, on choisit les fonctions-test dans l'espace de Sobolev  $H_0^1(\Omega)$ . On cherche ainsi la solution du problème

$$T'_0 \in H^1(\Omega), \quad \int_{\Omega} \Delta T'_0 \varphi \, d\mathbf{x} = 0 \quad \forall \varphi \in H_0^1(\Omega).$$

En intégrant par parties :

$$0 = \int_{\Omega} \Delta T'_0 \cdot \varphi d\mathbf{x} = - \int_{\Omega} \nabla T'_0 \cdot \nabla \varphi d\mathbf{x} - \int_{\partial\Omega} \underbrace{\frac{\partial T'_0}{\partial n}}_{=0} \varphi d\mathbf{x}$$

soit

$$\int_{\Omega} \nabla T'_0 \cdot \nabla \varphi d\mathbf{x} = 0 \quad \forall \varphi \in H_0^1(\Omega). \quad (3.5)$$

On se ramène à un problème homogène (c'est-à-dire à une solution cherchée dans l'espace  $H_0^1(\Omega)$ ) par un relèvement de la condition au bord. En supposant  $\theta_w \in H^{1/2}(\Gamma)$ , on définit  $R_\theta \in H^1(\Omega)$  un relèvement de  $\theta_w$  et on pose  $\tilde{T} = T'_0 - R_\theta$ . On a alors  $\tilde{T}|_\Gamma = 0$ , soit  $\tilde{T} \in H_0^1(\Omega)$  et le problème s'exprime en tenant compte de (3.5) : trouver  $\tilde{T} \in H_0^1(\Omega)$  tel que

$$\int_{\Omega} \nabla \tilde{T} \cdot \nabla \varphi d\mathbf{x} = - \int_{\Omega} \nabla R_\theta \cdot \nabla \varphi d\mathbf{x} \quad \forall \varphi \in H_0^1(\Omega). \quad (3.6)$$

La forme bilinéaire  $a : (\tilde{T}, \varphi) \in H_0^1(\Omega) \times H_0^1(\Omega) \mapsto \int_{\Omega} \nabla \tilde{T} \cdot \nabla \varphi d\mathbf{x}$  est

➤ continue :

$$\begin{aligned} |a(\tilde{T}, \varphi)| &= \left| \int_{\Omega} \nabla \tilde{T} \cdot \nabla \varphi d\mathbf{x} \right| \leq \int_{\Omega} |\nabla \tilde{T} \cdot \nabla \varphi| d\mathbf{x} \\ &\leq \left( \int_{\Omega} |\nabla \tilde{T}|^2 d\mathbf{x} \right)^{1/2} \left( \int_{\Omega} |\nabla \varphi|^2 d\mathbf{x} \right)^{1/2} \\ &= \|\tilde{T}\|_{H_0^1} \|\varphi\|_{H_0^1}, \end{aligned}$$

➤ coercive :  $a(\varphi, \varphi) = \int_{\Omega} |\nabla \varphi|^2 d\mathbf{x} = \|\nabla \varphi\|_{L^2}^2 = \|\varphi\|_{H_0^1}^2$ ,

et la forme linéaire  $L : \varphi \in H_0^1(\Omega) \mapsto - \int_{\Omega} \nabla R_\theta \cdot \nabla \varphi d\mathbf{x}$  est continue :

$$|L(\varphi)| = \left| \int_{\Omega} \nabla R_\theta \cdot \nabla \varphi d\mathbf{x} \right| \leq \|\nabla R_\theta\|_{L^2} \|\varphi\|_{H_0^1}.$$

Selon le théorème de Lax-Milgram, le problème (3.6) a donc une unique solution dans  $H_0^1$ . On a alors  $T'_0 = \tilde{T} + R_\theta$ , solution du problème initial.  $T'_0$  est également unique : si  $T_0^1$  et  $T_0^2$  sont deux solutions, alors  $T_0^1 - T_0^2 \in H_0^1(\Omega)$  est solution du problème

$$\begin{cases} \Delta \bar{T} = 0, \\ \bar{T}|_\Gamma = 0, \end{cases}$$

dont la solution triviale est  $\bar{T} \equiv 0$ . Le problème est donc bien posé.

### 3.2.2 Problème de Boussinesq : champ $\rho'_0$

Le problème de Boussinesq donnant le champ  $\rho'_0$  s'écrit

$$\begin{cases} \nabla (\rho'_0 + T'_0) = \mathbf{0}, \\ \rho'_{0\Gamma} = \sigma_w(f_0). \end{cases} \quad (3.7)$$



Le champ  $T'_0$  est désormais déterminé (jusqu'au bord du domaine de calcul) avec le problème de Laplace étudié au paragraphe 3.2.1, la seule inconnue du système (3.7) est donc  $\rho'_0$ .

Ce problème n'est pas de type Dirichlet : compte-tenu de l'expression de  $f_0$  et des conditions aux limites obtenues sur  $T'_0$  et  $\mathbf{u}_0'$ , la condition  $\rho'_{0\Gamma} = \sigma_w(f_0)$  n'est plus une condition aux limites, mais une condition de compatibilité. En effet, on a par définition de  $\sigma_w$  et avec les conditions d'adhérence (3.3)

$$\sigma_w(f_0) = -\frac{1}{2}\theta_w - \sqrt{2\pi} \langle (\mathbf{v} \cdot \mathbf{n})^- f_0 \rangle = \rho'_{0\Gamma} + \frac{1}{2}(\theta_w - T'_{0\Gamma}) - \frac{\sqrt{2\pi}}{2} \mathbf{u}_0'_{\Gamma} \cdot \mathbf{n} = \rho'_{0\Gamma}. \quad (3.8)$$

En remarquant que la quantité  $\rho' + T' = \frac{1}{3} \langle |\mathbf{v}|^2 f \rangle$  est homogène à une pression, on note

$$p' := \rho' + T', \quad \text{soit} \quad p'_n := \rho'_n + T'_n \quad \text{à chaque ordre du développement de Hilbert.}$$

Le problème (3.7) s'écrit alors simplement

$$\nabla p'_0 = \mathbf{0}. \quad (3.9)$$

La pression  $p'_0$  est donc une constante indéterminée. Ceci traduit un degré de liberté dans le problème BGK linéarisé : l'amplitude de la perturbation  $f$  n'est en effet pas fixée. On impose alors pour fermer le problème la condition

$$\int_{\Omega} p' d\mathbf{x} = 0,$$

qui donne à chaque ordre du développement de Hilbert

$$\boxed{\forall n \in \mathbb{N}, \quad \int_{\Omega} p'_n d\mathbf{x} = \int_{\Omega} (\rho'_n + T'_n) d\mathbf{x} = 0.} \quad (3.10)$$

La solution triviale, unique, du système (3.9)/(3.10) est  $p'_0 \equiv 0$ , soit

$$\boxed{\rho'_0 = -T'_0.}$$

### 3.2.3 Problème de Stokes sous contrainte : champ $\mathbf{u}_0'$

La vitesse  $\mathbf{u}_0'$  est quant à elle obtenue avec le problème de type Stokes "point-selle"

$$\begin{cases} \Delta(\mathbf{u}_0') - \nabla(\rho'_1 + T'_1) = 0, \\ \nabla \cdot \mathbf{u}_0' = 0, \\ \mathbf{u}_0'_{\Gamma} = \mathbf{u}_w. \end{cases}$$

☞ NB : la condition  $\mathbf{u}_w \cdot \mathbf{n} = 0$  imposée à la perturbation de vitesse à la paroi permet de vérifier à peu de frais la condition de compatibilité indispensable pour que le problème de Stokes soit bien posé. En effet, en intégrant la contrainte d'incompressibilité sur le domaine  $\Omega$ , on obtient avec la condition d'adhérence

$$0 = \int_{\Omega} \nabla \cdot \mathbf{u}_0' d\mathbf{x} = \int_{\Gamma} \mathbf{u}_0' \cdot \mathbf{n} dl = \int_{\Gamma} \mathbf{u}_w \cdot \mathbf{n} dl,$$

soit une condition "faible" sur la vitesse normale à la paroi.  $\square$

La condition aux limites impose de nouveau de rechercher une solution en vitesse dans l'espace  $H_0^1(\Omega, \mathbb{R}^3)$  après un relèvement ; les deux termes de la première équation du système s'écrivent donc sous la forme faible

$$\begin{cases} \int_{\Omega} \Delta(\mathbf{u}_0') \cdot \boldsymbol{\varphi} d\mathbf{x} = - \int_{\Omega} \nabla \mathbf{u}_0' : \nabla \boldsymbol{\varphi} d\mathbf{x} - \int_{\partial\Omega} \underbrace{\nabla \mathbf{u}_0' : (\boldsymbol{\varphi} \otimes \mathbf{n})}_{=0} d\mathbf{x} \quad \forall \boldsymbol{\varphi} \in H_0^1(\Omega, \mathbb{R}^3), \\ \int_{\Omega} \nabla (\rho_1' + T_1') \cdot \boldsymbol{\varphi} d\mathbf{x} = - \int_{\Omega} (\rho_1' + T_1') \nabla \cdot \boldsymbol{\varphi} d\mathbf{x} - \int_{\partial\Omega} (\rho_1' + T_1') \underbrace{\boldsymbol{\varphi} \cdot \mathbf{n}}_{=0} d\mathbf{x} \quad \forall \boldsymbol{\varphi} \in H_0^1(\Omega, \mathbb{R}^3). \end{cases}$$

Comme pour le problème de Boussinesq, la quantité  $p_1' = \rho_1' + T_1'$  doit être recherchée dans l'espace

$$L_0^2(\Omega) = \left\{ \psi \in L^2(\Omega), \int_{\Omega} \psi d\mathbf{x} = 0 \right\},$$

conformément à la condition (3.10). On obtient ainsi le problème variationnel sous contrainte : trouver  $(\mathbf{u}_0', p_1) \in H^1(\Omega, \mathbb{R}^3) \times L_0^2$  solution de

$$\begin{cases} \int_{\Omega} \nabla \mathbf{u}_0' : \nabla \boldsymbol{\varphi} d\mathbf{x} - \int_{\Omega} p_1' \nabla \cdot \boldsymbol{\varphi} d\mathbf{x} = 0 \quad \forall \boldsymbol{\varphi} \in H_0^1(\Omega, \mathbb{R}^3), \\ \int_{\Omega} \nabla \cdot \mathbf{u}_0' \psi d\mathbf{x} = 0 \quad \forall \psi \in L_0^2(\Omega), \\ \mathbf{u}_0'|_{\Gamma} = \mathbf{u}_w. \end{cases} \quad (3.11)$$

Comme pour le problème de Laplace, on effectue un relèvement  $\mathbf{R}_u \in H^1(\Omega, \mathbb{R}^3)$  de la vitesse, et on cherche la solution  $(\tilde{\mathbf{u}}, p_1') \in H_0^1(\Omega, \mathbb{R}^3) \times L_0^2$  de :

$$\begin{cases} \int_{\Omega} \nabla \tilde{\mathbf{u}} : \nabla \boldsymbol{\varphi} d\mathbf{x} - \int_{\Omega} p_1' \nabla \cdot \boldsymbol{\varphi} d\mathbf{x} = - \int_{\Omega} \nabla \cdot \mathbf{R}_u \nabla \cdot \boldsymbol{\varphi} d\mathbf{x} \quad \forall \boldsymbol{\varphi} \in H_0^1(\Omega, \mathbb{R}^3), \\ \int_{\Omega} \nabla \cdot \tilde{\mathbf{u}} \psi d\mathbf{x} = - \int_{\Omega} \nabla \cdot \mathbf{R}_u \psi d\mathbf{x} \quad \forall \psi \in L_0^2(\Omega). \end{cases} \quad (3.12)$$

que l'on écrit

$$\begin{cases} a(\tilde{\mathbf{u}}, \boldsymbol{\varphi}) + b(p_1', \boldsymbol{\varphi}) = L(\boldsymbol{\varphi}) \quad \forall \boldsymbol{\varphi} \in H_0^1(\Omega, \mathbb{R}^3), \\ -b(\tilde{\mathbf{u}}, \psi) = G(\psi) \quad \forall \psi \in L_0^2(\Omega). \end{cases} \quad (3.13)$$

avec

$$\begin{cases} a : (\mathbf{u}, \boldsymbol{\varphi}) \in H_0^1(\Omega, \mathbb{R}^3) \times H_0^1(\Omega, \mathbb{R}^3) \mapsto \int_{\Omega} \nabla \mathbf{u} : \nabla \boldsymbol{\varphi} d\mathbf{x}, \\ b : (\boldsymbol{\varphi}, p) \in H_0^1(\Omega, \mathbb{R}^3) \times L_0^2(\Omega) \mapsto - \int_{\Omega} p \nabla \cdot \boldsymbol{\varphi} d\mathbf{x}, \\ L : \boldsymbol{\varphi} \in H_0^1(\Omega, \mathbb{R}^3) \mapsto - \int_{\Omega} \nabla \cdot \mathbf{R}_u \nabla \cdot \boldsymbol{\varphi} d\mathbf{x}, \\ G : \psi \in L^2 \mapsto \int_{\Omega} \nabla \cdot \mathbf{R}_u \psi d\mathbf{x}. \end{cases}$$

On montre l'existence et l'unicité de cette solution en utilisant le théorème de Lax-Milgram généralisé. On va utiliser pour cela le résultat suivant :

**Proposition 1.**

$$\exists D > 0, \forall \mathbf{v} \in H_0^1(\Omega, \mathbb{R}^3), \|\nabla \cdot \mathbf{v}\|_{L^2} \leq D \|\mathbf{v}\|_{H_0^1}.$$

*Démonstration 1.* Soit  $\mathbf{v} \in H_0^1(\Omega, \mathbb{R}^3)$ . On pose  $\mathbf{J} = \left( \frac{\partial v_i}{\partial x_i} \right)_i$ . On a alors

$$\begin{aligned} \left| \sum_i \frac{\partial v_i}{\partial x_i} \right| &\leq \sum_i \left| \frac{\partial v_i}{\partial x_i} \right| = \|\mathbf{J}\|_{L^1} \leq D \|\mathbf{J}\|_{L^2}, \quad D > 0 \quad (\text{équivalence des normes sur } \mathbb{R}^3) \\ &= D \left[ \sum_i \left( \frac{\partial v_i}{\partial x_i} \right)^2 \right]^{1/2} \\ &\leq D \left[ \sum_{i,j} \left( \frac{\partial v_i}{\partial x_j} \right)^2 \right]^{1/2}. \end{aligned}$$

On obtient donc  $\int_{\Omega} \left( \sum_i \frac{\partial v_i}{\partial x_i} \right)^2 d\mathbf{x} \leq D^2 \sum_{i,j} \int_{\Omega} \left( \frac{\partial v_i}{\partial x_j} \right)^2 d\mathbf{x}$ , soit  $\|\nabla \cdot \mathbf{v}\|_{L^2} \leq D \|\mathbf{v}\|_{H_0^1}$ .  $\square$

- La forme bilinéaire  $a$  est

➤ continue :

$$\begin{aligned} |a(\mathbf{u}, \boldsymbol{\varphi})| &= \left| \int_{\Omega} \nabla \mathbf{u} : \nabla \boldsymbol{\varphi} d\mathbf{x} \right| = \left| \sum_i \int_{\Omega} \nabla \mathbf{u}_i \cdot \nabla \boldsymbol{\varphi}_i d\mathbf{x} \right| \leq \sum_i \left| \int_{\Omega} \nabla \mathbf{u}_i \cdot \nabla \boldsymbol{\varphi}_i d\mathbf{x} \right| \\ &\leq \sum_i \|\mathbf{u}_i\|_{H_0^1} \|\boldsymbol{\varphi}_i\|_{H_0^1} \\ &\leq \|\mathbf{u}\|_{H_0^1} \|\boldsymbol{\varphi}\|_{H_0^1} \end{aligned}$$

(selon l'inégalité de Cauchy-Schwarz),

➤ coercive :  $a(\boldsymbol{\varphi}, \boldsymbol{\varphi}) = \sum_i \int_{\Omega} |\nabla \boldsymbol{\varphi}_i|_{L^2}^2 d\mathbf{x} = \sum_i \|\nabla \boldsymbol{\varphi}_i\|_{L^2}^2 = \|\boldsymbol{\varphi}\|_{H_0^1}^2$ ;

- La forme bilinéaire  $b$

➤ est continue :

$$\begin{aligned} |b(\boldsymbol{\varphi}, \rho)| &= \left| \int_{\Omega} \rho \nabla \cdot \boldsymbol{\varphi} d\mathbf{x} \right| \leq \int_{\Omega} |\rho \nabla \cdot \boldsymbol{\varphi}| d\mathbf{x} \leq \left( \int_{\Omega} |\rho|^2 \right)^{1/2} \left( \int_{\Omega} |\nabla \cdot \boldsymbol{\varphi}|^2 \right)^{1/2} \\ &= \|\rho\|_{L^2} \|\nabla \cdot \boldsymbol{\varphi}\|_{L^2} \leq D \|\rho\|_{L^2} \|\boldsymbol{\varphi}\|_{H_0^1} \end{aligned}$$

selon la proposition 1.

➤ vérifie la condition inf-sup suivante :

$$\forall \rho \in L_0^2, \sup_{\boldsymbol{\varphi} \in H_0^1(\Omega, \mathbb{R}^3) \setminus \{\mathbf{0}\}} \frac{b(\boldsymbol{\varphi}, \rho)}{\|\boldsymbol{\varphi}\|_{H_0^1}} \geq \beta \|\rho\|_{L^2}, \quad \beta > 0. \quad (3.14)$$

En effet, d'après le lemme de Nečas :

$$\exists C > 0, \forall \rho \in L_0^2(\Omega), \exists \boldsymbol{\psi} \in H_0^1(\Omega, \mathbb{R}^3) / \begin{cases} \rho = -\nabla \cdot \boldsymbol{\psi}, \\ \|\boldsymbol{\psi}\|_{H_0^1} \leq C \|\rho\|_{L^2}. \end{cases} \quad (3.15)$$

Ainsi, pour tout  $\rho \in L_0^2(\Omega)$ ,  $b(\boldsymbol{\psi}, \rho) = -\int_{\Omega} \rho \nabla \cdot \boldsymbol{\psi} \, d\mathbf{x} = \int_{\Omega} \rho^2 \, d\mathbf{x} = \|\rho\|_{L^2}^2$ , et

$$\frac{b(\boldsymbol{\psi}, \rho)}{\|\boldsymbol{\psi}\|_{H_0^1}} = \|\rho\|_{L^2}^2 \|\boldsymbol{\psi}\|_{H_0^1}^{-1} \geq C^{-1} \|\rho\|_{L^2}.$$

☞ NB : la condition précédente est dite "inf-sup" car elle peut également s'écrire

$$\inf_{\rho \in L_0^2(\Omega)} \sup_{\boldsymbol{\varphi} \in H_0^1(\Omega, \mathbb{R}^3) \setminus \{0\}} \frac{b(\boldsymbol{\varphi}, \rho)}{\|\boldsymbol{\varphi}\|_{H_0^1}} \geq \beta \|\rho\|_{L^2}, \quad \beta > 0.$$

On parle aussi de "condition de stabilité" pour la forme bilinéaire  $b$ .  $\square$

- Enfin, les formes linéaires  $L$  et  $G$  sont clairement continues.

D'après le théorème de Lax-Milgram généralisé, le système (3.12) a donc une unique solution, et on déduit comme pour le problème de Laplace que le problème (3.11) est lui aussi bien posé, de solution  $(\mathbf{u}_0', p_1') = (\tilde{\mathbf{u}} + \mathbf{R}_{\mathbf{u}}, p_1')$ .

☞ NB : le théorème de Lax-Milgram généralisé appliqué au problème (3.12) est équivalent au théorème de Banach-Nečas-Babuška appliqué au problème

$$c((\tilde{\mathbf{u}}, p_1'), (\boldsymbol{\varphi}, \psi)) = H(\boldsymbol{\varphi}, \psi) \quad \forall (\boldsymbol{\varphi}, \psi) \in H_0^1(\Omega, \mathbb{R}^3) \times L_0^2(\Omega)$$

avec  $c((\tilde{\mathbf{u}}, p_1'), (\boldsymbol{\varphi}, \psi)) = a(\tilde{\mathbf{u}}, \boldsymbol{\varphi}) + b(\boldsymbol{\varphi}, p_1') - b(\tilde{\mathbf{u}}, \psi)$  et  $H(\boldsymbol{\varphi}, \psi) = L(\boldsymbol{\varphi}) + G(\psi)$ ; ce problème est obtenu simplement en sommant les deux équations du système (3.12).  $\square$

Il existe un cas particulier important, celui où la donnée au bord  $\mathbf{u}_w$  est constante. Le problème (3.11) a alors pour solution triviale

$$(\mathbf{u}_0', p_1) \equiv (\mathbf{u}_w, 0).$$

Ce cas est courant dans la pratique, celui d'une unique paroi indéformable.

### 3.3 Écriture du schéma Galerkin Discontinu

L'étude du schéma Galerkin Discontinu appliqué au problème BGK linéarisé effectuée dans les parties suivantes est inspirée des travaux de Jean-Luc Guermond et Guido Kanschat sur l'équation de transport linéaire dans [37].

### 3.3.1 Forme faible de l'équation BGK linéarisée

On construit le schéma Galerkin Discontinu via une formulation faible du problème. L'espace fonctionnel où on recherche les distributions est déduit naturellement des propriétés de l'opérateur BGK linéarisé  $L$  sur l'espace des vitesses.

On rappelle, ce qui a été prouvé au chapitre précédent, que l'opérateur  $L$  vérifie les lois de conservation

$$\forall g \in L^2(\Omega \times \mathbb{R}^3), \quad \begin{cases} \langle L(g) \rangle = 0, \\ \langle \mathbf{v} L(g) \rangle = \mathbf{0}, \\ \langle |\mathbf{v}|^2 L(g) \rangle = 0. \end{cases} \quad (3.16)$$

avec la notation usuelle  $\int_{\mathbb{R}^3} \cdot d\mathbf{v} = \langle \cdot \rangle$ .

La propriété essentielle de cette opérateur, qui va motiver le choix de l'espace de Hilbert dans la formulation faible du problème BGK linéarisé, est la suivante :

**Proposition 2.** L'opérateur  $L$  est autoadjoint sur l'espace  $L^2(\mathbb{R}^3, M^{-1}d\mathbf{v})$ .

*Démonstration 2.* On note

$$\begin{cases} \mathbf{m}(\mathbf{v}) = \left( 1, \mathbf{v}, \frac{|\mathbf{v}|^2 - 3}{3} \right)^T, \\ \boldsymbol{\alpha}(\mathbf{v}) = \left( 1, \mathbf{v}, \frac{|\mathbf{v}|^2 - 3}{2} \right)^T. \end{cases}$$

L'opérateur  $L$  s'exprime avec ces deux quantités

$$L(f) = \boldsymbol{\alpha} \cdot \langle \mathbf{m} f \rangle M - f.$$

On a alors

$$\begin{aligned} (L(f)|g) &:= \int_{\mathbb{R}^3} L(f) g M^{-1} d\mathbf{v} \\ &= \int_{\mathbb{R}^3} (\boldsymbol{\alpha} \cdot \langle \mathbf{m} f \rangle M - f) g M^{-1} d\mathbf{v} \\ &= \int_{\mathbb{R}^3} \boldsymbol{\alpha} \cdot \int_{\mathbb{R}^3} \mathbf{m}(\mathbf{w}) f d\mathbf{w} g d\mathbf{v} - \int_{\mathbb{R}^3} f g M^{-1} \int_{\mathbb{R}^3} M(\mathbf{w}) d\mathbf{w} d\mathbf{v} \quad \text{car } \int_{\mathbb{R}^3} M(\mathbf{w}) d\mathbf{w} = 1 \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \boldsymbol{\alpha}(\mathbf{v}) \cdot \mathbf{m}(\mathbf{w}) f(\mathbf{w}) g(\mathbf{v}) d\mathbf{w} d\mathbf{v} - \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f(\mathbf{v}) g(\mathbf{v}) M^{-1}(\mathbf{v}) M(\mathbf{w}) d\mathbf{w} d\mathbf{v}, \end{aligned}$$

or  $\boldsymbol{\alpha}(\mathbf{v}) \cdot \mathbf{m}(\mathbf{w}) = \boldsymbol{\alpha}(\mathbf{w}) \cdot \mathbf{m}(\mathbf{v})$ , donc

$$\begin{aligned} (L(f)|g) &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \boldsymbol{\alpha}(\mathbf{w}) \cdot \mathbf{m}(\mathbf{v}) f(\mathbf{w}) g(\mathbf{v}) d\mathbf{w} d\mathbf{v} - \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f(\mathbf{w}) g(\mathbf{w}) M^{-1}(\mathbf{w}) M(\mathbf{v}) d\mathbf{w} d\mathbf{v} \\ &= \int_{\mathbb{R}^3} \boldsymbol{\alpha} \cdot \langle \mathbf{m} g \rangle f d\mathbf{w} - \int_{\mathbb{R}^3} f g M^{-1} d\mathbf{w} \\ &= (L(g)|f). \quad \square \end{aligned}$$

On va donc considérer les fonctions de distribution dans l'espace

$$\mathbb{E} = L^2(\Omega \times \mathbb{R}^3, M^{-1} d\mathbf{x} d\mathbf{v}).$$

On écrit alors la formulation faible de l'équation (3.1) :

$$\forall g \in \mathbb{E}, \int_{\Omega \times \mathbb{R}^3} \mathbf{v} \cdot \nabla_x f g M^{-1} d\mathbf{x} d\mathbf{v} - \frac{1}{\varepsilon} \int_{\Omega \times \mathbb{R}^3} L(f) g M^{-1} d\mathbf{x} d\mathbf{v} = 0,$$

soit

$$\int_{\Omega} \langle g \mathbf{v} \cdot \nabla_x f M^{-1} \rangle d\mathbf{x} - \frac{1}{\varepsilon} \int_{\Omega} \langle L(f) g M^{-1} \rangle d\mathbf{x} = 0. \quad (3.17)$$

NB : on se place ici dans un cadre continu en vitesses. L'étude des schémas Galerkin Discontinu à vitesses discrètes demanderait un calcul soigneux de la maxwellienne linéarisée discrète, dont l'expression est a priori différente de celle de la maxwellienne linéarisée continue.  $\square$

### 3.3.2 Approximation Galerkin Discontinu

#### 3.3.2.1 Définitions et notations

On considère une partition  $\mathbb{T}$  du domaine  $\Omega$

$$\mathbb{T} := \left\{ C, \Omega \cup \Gamma = \bigcup \overline{C} \right\},$$

et deux espaces d'approximation  $\mathbb{X}_k$  et  $\mathbb{Y}_k$  de fonctions polynomiales par morceaux, respectivement scalaires et vectorielles, de degré  $k$  sur chaque maille de  $\mathbb{T}$  :

$$\mathbb{X}_k := \left\{ g \in L^2(\Omega), g|_C \in \mathbb{P}_k, \forall C \in \mathbb{T} \right\}, \quad \mathbb{Y}_k := \left\{ g \in L^2(\Omega, \mathbb{R}^3), g|_C \in (\mathbb{P}_k)^3, \forall C \in \mathbb{T} \right\}.$$

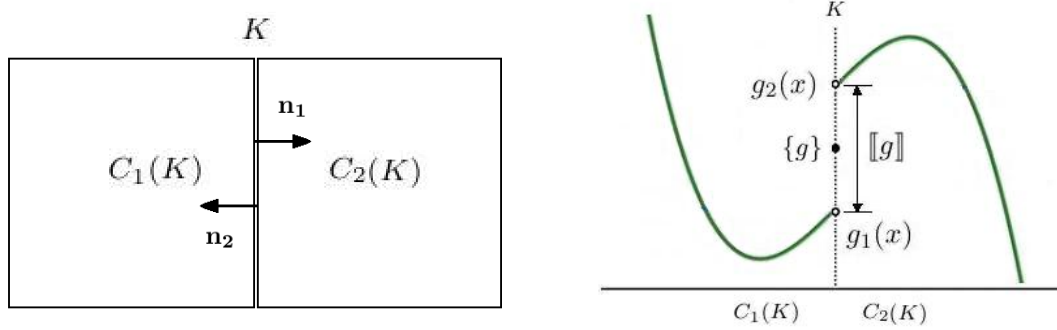
On introduit alors l'espace d'approximation pour les fonctions de distribution, soit un espace d'approximation de  $\mathbb{E} = L^2(\Omega \times \mathbb{R}^3, M^{-1} d\mathbf{x} d\mathbf{v})$

$$\mathbb{E}_k := \left\{ f \in \mathbb{E} / \mathbf{x} \mapsto f(\mathbf{x}, \mathbf{v}) \in \mathbb{X}_k \forall \mathbf{v} \in \mathbb{R}^3 \text{ et } \mathbf{v} \mapsto f(\mathbf{x}, \mathbf{v}) \in L^2(\mathbb{R}^3, M^{-1} d\mathbf{v}) \right\}$$

On note  $\mathbb{F} := \left\{ \overline{C} \setminus \overset{\circ}{C}, C \in \mathbb{T} \right\}$  l'ensemble des arêtes de  $\mathbb{T}$ . L'ensemble  $\mathbb{F}$  est séparé en un ensemble d'arêtes internes au domaine de calcul, noté  $\overset{\circ}{\mathbb{F}}$ , et un ensemble d'arêtes de bord, noté  $\mathbb{F}^\partial$  :

$$\begin{cases} \overset{\circ}{\mathbb{F}} := \{ K \in \mathbb{F}, K \cap \Gamma = \emptyset \}, \\ \mathbb{F}^\partial := \mathbb{F} \setminus \overset{\circ}{\mathbb{F}}. \end{cases}$$

Pour toute arête  $K \in \overset{\circ}{\mathbb{F}}$ , on note  $C_1(K)$  et  $C_2(K)$  les deux mailles adjacentes à cette arête, de normales sortantes  $\mathbf{n}_1$  et  $\mathbf{n}_2$  respectivement, comme indiqué sur la figure 3.2.

FIGURE 3.2: Normales et discontinuités de la fonction  $g$  aux arêtes du maillage

On définit alors pour toute fonction  $g \in \mathbb{E}$  et tout point  $x \in K$  les limites de  $g$  de chaque côté de l'arête ainsi que le saut et la valeur moyenne de  $g$  à travers cette arête :

$$\begin{cases} g_1(x) := \lim_{\substack{y \rightarrow x \\ y \in C_1(K)}} g(y), \\ g_2(x) := \lim_{\substack{y \rightarrow x \\ y \in C_2(K)}} g(y), \\ [g] := g_2 - g_1, \\ \{g\} := \frac{1}{2} (g_1 + g_2). \end{cases}$$

Ces définitions étant posées, on va maintenant écrire une formulation générale du schéma Galerkin Discontinu appliqué au problème BGK linéarisé (3.1).

### 3.3.2.2 Construction du schéma Galerkin Discontinu

L'équation (3.17) s'écrit sur le maillage  $\mathbb{T}$

$$\sum_{C \in \mathbb{T}} \int_C \langle g \mathbf{v} \cdot \nabla_x f M^{-1} \rangle d\mathbf{x} - \frac{1}{\varepsilon} \sum_{C \in \mathbb{T}} \int_C \langle L(f) g M^{-1} \rangle d\mathbf{x} = 0.$$

**Proposition 3.** Pour tous  $\gamma = \beta M \in \mathbb{E}_k$  et  $\alpha \in \mathbb{E}$  continue par rapport à  $\mathbf{x}$  telle que  $\alpha(x \in \Gamma, \mathbf{v}) = \alpha|_{\Gamma}$  si  $\mathbf{v} \cdot \mathbf{n} > 0$  :

$$\begin{aligned} \sum_{C \in \mathbb{T}} \int_C \langle \gamma \mathbf{v} \cdot \nabla_x \alpha M^{-1} \rangle d\mathbf{x} &= \sum_{C \in \mathbb{T}} \int_C \langle \beta \mathbf{v} \cdot \nabla_x \alpha \rangle d\mathbf{x} \\ &= - \sum_{C \in \mathbb{T}} \int_C \langle \alpha \mathbf{v} \cdot \nabla_x \beta \rangle d\mathbf{x} + \sum_{K \in \mathring{\mathbb{F}}} \int_K \langle \mathbf{v} \cdot \tilde{\mathbf{n}} \alpha [\beta] \rangle d\mathbf{x} \\ &\quad - \sum_{K \in \mathbb{F}^{\partial}} \int_{K, \mathbf{v} \cdot \mathbf{n} < 0} \langle \mathbf{v} \cdot \mathbf{n} \alpha \beta \rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^{\partial}} \int_{K, \mathbf{v} \cdot \mathbf{n} > 0} \langle \mathbf{v} \cdot \mathbf{n} \alpha|_{\Gamma} \beta \rangle d\mathbf{x}. \end{aligned}$$

*Démonstration 3.* En intégrant l'équation par parties (formule de Green), on obtient

$$\sum_{C \in \mathbb{T}} \int_C \langle \beta \mathbf{v} \cdot \nabla_x \alpha \rangle d\mathbf{x} = - \sum_{C \in \mathbb{T}} \int_C \langle \alpha \mathbf{v} \cdot \nabla_x \beta \rangle d\mathbf{x} + \sum_{C \in \mathbb{T}} \int_{\partial C} \langle \mathbf{v} \cdot \mathbf{n}_s \alpha \beta \rangle d\mathbf{x},$$

où  $\mathbf{n}_s$  est la normale sortante à la maille  $C$ .

Les termes de bords se répartissent sur les arêtes internes et les arêtes de bord :

$$\sum_{C \in \mathbb{T}} \int_{\partial C} \langle \mathbf{v} \cdot \mathbf{n}_s \alpha \beta \rangle d\mathbf{x} = \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n}_s \alpha \beta)_2 + (\mathbf{v} \cdot \mathbf{n}_s \alpha \beta)_1 \right\rangle d\mathbf{x} + \sum_{K \in \mathbb{F}^\partial} \int_K \langle \mathbf{v} \cdot \mathbf{n}_s \alpha \beta \rangle d\mathbf{x}.$$

En choisissant pour les arêtes internes la normale  $\mathbf{n}_2$  comme normale de référence, notée  $\tilde{\mathbf{n}}$ , on obtient

$$\begin{aligned} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n}_s \alpha \beta)_2 + (\mathbf{v} \cdot \mathbf{n}_s \alpha \beta)_1 \right\rangle d\mathbf{x} &= \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle (\mathbf{v} \cdot \tilde{\mathbf{n}} \alpha_2 \beta_2 - \mathbf{v} \cdot \tilde{\mathbf{n}} \alpha_1 \beta_1) \right\rangle d\mathbf{x} \\ &= \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [\alpha \beta] \right\rangle d\mathbf{x}. \end{aligned}$$

La condition aux limites appliquée sur la fonction  $\alpha$  au niveau des arêtes externes dépend du signe du produit scalaire  $\mathbf{v} \cdot \mathbf{n}_s$ . On décompose donc les termes de bord, en les exprimant avec la normale  $\mathbf{n}$  dirigée vers le domaine fluide ( $\mathbf{n} = -\mathbf{n}_s$ ) :

$$\begin{aligned} \sum_{K \in \mathbb{F}^\partial} \int_K \langle \mathbf{v} \cdot \mathbf{n}_s \alpha \beta \rangle d\mathbf{x} &= - \sum_{K \in \mathbb{F}^\partial} \int_K \langle \mathbf{v} \cdot \mathbf{n} \alpha \beta \rangle d\mathbf{x} \\ &= - \sum_{K \in \mathbb{F}^\partial} \int_{K, \mathbf{v} \cdot \mathbf{n} < 0} \langle \mathbf{v} \cdot \mathbf{n} \alpha \beta \rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_{K, \mathbf{v} \cdot \mathbf{n} > 0} \langle \mathbf{v} \cdot \mathbf{n} \alpha \beta \rangle d\mathbf{x} \\ &= - \sum_{K \in \mathbb{F}^\partial} \int_{K, \mathbf{v} \cdot \mathbf{n} < 0} \langle \mathbf{v} \cdot \mathbf{n} \alpha \beta \rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_{K, \mathbf{v} \cdot \mathbf{n} > 0} \langle \mathbf{v} \cdot \mathbf{n} \alpha|_\Gamma \beta \rangle d\mathbf{x}, \end{aligned}$$

ce qui donne donc

$$\begin{aligned} \sum_{C \in \mathbb{T}} \int_C \langle \beta \mathbf{v} \cdot \nabla_x \alpha \rangle d\mathbf{x} &= - \sum_{C \in \mathbb{T}} \int_C \langle \alpha \mathbf{v} \cdot \nabla_x \beta \rangle d\mathbf{x} + \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [\alpha \beta] \right\rangle d\mathbf{x} \\ &\quad - \sum_{K \in \mathbb{F}^\partial} \int_{K, \mathbf{v} \cdot \mathbf{n} < 0} \langle \mathbf{v} \cdot \mathbf{n} \alpha \beta \rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_{K, \mathbf{v} \cdot \mathbf{n} > 0} \langle \mathbf{v} \cdot \mathbf{n} \alpha|_\Gamma \beta \rangle d\mathbf{x}, \end{aligned} \tag{3.18}$$

et par continuité de  $\alpha : \mathbf{v} \cdot \tilde{\mathbf{n}} [\alpha \beta] = \mathbf{v} \cdot \tilde{\mathbf{n}} \alpha [\beta]$ , d'où le résultat voulu.  $\square$

En appliquant ce résultat au problème (3.1), on obtient

$$\begin{aligned} \sum_{C \in \mathbb{T}} \int_C \left\langle f \mathbf{v} \cdot \nabla_x g M^{-1} + \frac{1}{\varepsilon} L(f) g M^{-1} \right\rangle d\mathbf{x} &- \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} f [g] M^{-1} \right\rangle d\mathbf{x} \\ &+ \sum_{K \in \mathbb{F}^\partial} \int_{K, \mathbf{v} \cdot \mathbf{n} < 0} \langle \mathbf{v} \cdot \mathbf{n} f g M^{-1} \rangle d\mathbf{x} = - \sum_{K \in \mathbb{F}^\partial} \int_{K, \mathbf{v} \cdot \mathbf{n} > 0} \langle \mathbf{v} \cdot \mathbf{n} p_w(f) g \rangle d\mathbf{x} \quad \forall g \in \mathbb{E}_k. \end{aligned}$$



On approche ensuite la distribution continue  $f$  par une distribution discrète  $f^h \in \mathbb{E}_k$ . On doit alors définir le terme de flux sur les arêtes internes. Pour  $\varphi \in \mathbb{E}_k$ , on définit les flux "upwind" et "downwind" aux interfaces par :

$$\begin{cases} \mathbf{v} \cdot \tilde{\mathbf{n}} \lfloor \varphi \rfloor := (\mathbf{v} \cdot \tilde{\mathbf{n}})^+ \varphi_2 + (\mathbf{v} \cdot \tilde{\mathbf{n}})^- \varphi_1, \\ \mathbf{v} \cdot \tilde{\mathbf{n}} \lceil \varphi \rceil := (\mathbf{v} \cdot \tilde{\mathbf{n}})^- \varphi_2 + (\mathbf{v} \cdot \tilde{\mathbf{n}})^+ \varphi_1, \end{cases}$$

où

$$\forall a \in \mathbb{R}, \quad a^\pm = a \text{ si } a \gtrless 0 \text{ et } 0 \text{ sinon.}$$

**On choisit une approximation upwind** : on cherche  $f^h \in \mathbb{E}_k$  telle que

$$\begin{aligned} \sum_{C \in \mathbb{T}} \int_C \left\langle f^h \mathbf{v} \cdot \nabla_x g M^{-1} + \frac{1}{\varepsilon} L(f^h) g M^{-1} \right\rangle d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} \lfloor f^h \rfloor \llbracket g \rrbracket M^{-1} \right\rangle d\mathbf{x} \\ + \sum_{K \in \mathbb{F}^\partial} \int_K \langle (\mathbf{v} \cdot \mathbf{n})^- f^h g M^{-1} \rangle d\mathbf{x} = - \sum_{K \in \mathbb{F}^\partial} \int_K \langle (\mathbf{v} \cdot \mathbf{n})^+ p_w(f^h) g \rangle d\mathbf{x} \quad \forall g \in \mathbb{E}_k. \end{aligned} \quad (3.19)$$

On peut écrire cette équation sous une autre forme faisant intervenir un flux downwind sur la fonction-test, et qui va être utile dans la suite. On note d'abord la propriété

$$\forall \alpha, \beta \in \mathbb{E}_k, \quad \llbracket \alpha \beta \rrbracket = \lfloor \alpha \rfloor \llbracket \beta \rrbracket + \lceil \beta \rceil \llbracket \alpha \rrbracket. \quad (3.20)$$

En intégrant une nouvelle fois par parties le premier terme de l'équation et en utilisant la relation (3.20), on obtient

$$\begin{aligned} \sum_{C \in \mathbb{T}} \int_C \left\langle f^h \mathbf{v} \cdot \nabla_x g M^{-1} \right\rangle d\mathbf{x} &= - \sum_{C \in \mathbb{T}} \int_C \left\langle g \mathbf{v} \cdot \nabla_x f^h M^{-1} \right\rangle d\mathbf{x} + \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} \llbracket f^h g \rrbracket M^{-1} \right\rangle d\mathbf{x} \\ &\quad - \sum_{K \in \mathbb{F}^\partial} \int_K \langle (\mathbf{v} \cdot \mathbf{n})^- f^h g M^{-1} \rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \langle (\mathbf{v} \cdot \mathbf{n})^+ f^h g M^{-1} \rangle d\mathbf{x} \\ &= - \sum_{C \in \mathbb{T}} \int_C \left\langle g \mathbf{v} \cdot \nabla_x f^h M^{-1} \right\rangle d\mathbf{x} + \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} \lfloor f^h \rfloor \llbracket g \rrbracket M^{-1} + \mathbf{v} \cdot \tilde{\mathbf{n}} \lceil g \rceil \llbracket f^h \rrbracket M^{-1} \right\rangle d\mathbf{x} \\ &\quad - \sum_{K \in \mathbb{F}^\partial} \int_K \langle (\mathbf{v} \cdot \mathbf{n})^- f^h g M^{-1} \rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \langle (\mathbf{v} \cdot \mathbf{n})^+ f^h g M^{-1} \rangle d\mathbf{x}, \end{aligned}$$

ce qui donne

$$\begin{aligned} \sum_{C \in \mathbb{T}} \int_C \left\langle g \mathbf{v} \cdot \nabla_x f^h M^{-1} - \frac{1}{\varepsilon} L(f^h) g M^{-1} \right\rangle d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} \lceil g \rceil \llbracket f^h \rrbracket M^{-1} \right\rangle d\mathbf{x} \\ + \sum_{K \in \mathbb{F}^\partial} \int_K \langle (\mathbf{v} \cdot \mathbf{n})^+ f^h g M^{-1} \rangle d\mathbf{x} = \sum_{K \in \mathbb{F}^\partial} \int_K \langle (\mathbf{v} \cdot \mathbf{n})^+ p_w(f^h) g \rangle d\mathbf{x} \quad \forall g \in \mathbb{E}_k. \end{aligned} \quad (3.21)$$

Le schéma Galerkin Discontinu étant défini, on va pouvoir procéder à son analyse asymptotique d'une façon similaire à ce qui a été fait précédemment sur le problème BGK linéarisé continu, c'est-à-dire en effectuant un développement de Hilbert sur la distribution discrète  $f^h$ , puis en séparant les termes en fonction des puissances de  $\varepsilon$ .

### 3.4 Analyse asymptotique du schéma Galerkin Discontinu

#### 3.4.1 Développement de Hilbert

On effectue de la même façon que sur l'équation continue un développement de Hilbert sur la distribution  $f^h$  :

$$f^h(\mathbf{x}, \mathbf{v}, \varepsilon) = f_0^h(\mathbf{x}, \mathbf{v}) + \varepsilon f_1^h(\mathbf{x}, \mathbf{v}) + \dots + \varepsilon^n f_n^h(\mathbf{x}, \mathbf{v}) + \mathcal{O}(\varepsilon^{n+1})$$

Ce développement introduit dans l'équation (3.19), les différents termes s'écrivent

$$\begin{aligned} & \sum_{C \in \mathbb{T}} \int_C \left\langle f^h \mathbf{v} \cdot \nabla_x g M^{-1} + \frac{1}{\varepsilon} L(f^h) g M^{-1} \right\rangle d\mathbf{x} = \varepsilon^{-1} \sum_{C \in \mathbb{T}} \int_C \left\langle L(f_0^h) g M^{-1} \right\rangle d\mathbf{x} \\ & \quad + \sum_{i \geq 0} \varepsilon^i \sum_{C \in \mathbb{T}} \int_C \left\langle f_i^h \mathbf{v} \cdot \nabla_x g M^{-1} + L(f_{i+1}^h) g M^{-1} \right\rangle d\mathbf{x}, \\ & \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}}[f^h] \llbracket g \rrbracket M^{-1} \right\rangle d\mathbf{x} = \sum_{i \geq 0} \varepsilon^i \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}}[f_i^h] \llbracket g \rrbracket M^{-1} \right\rangle d\mathbf{x}, \\ & \sum_{K \in \mathbb{F}^\partial} \int_K \langle f^h g M^{-1} (\mathbf{v} \cdot \mathbf{n})^\pm \rangle d\mathbf{x} = \sum_{i \geq 0} \varepsilon^i \sum_{K \in \mathbb{F}^\partial} \int_K \langle f_i^h g M^{-1} (\mathbf{v} \cdot \mathbf{n})^\pm \rangle d\mathbf{x}, \\ & \sum_{K \in \mathbb{F}^\partial} \int_K \langle p_w(f^h) (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x} = \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle p_w \left( \sum_{i \geq 0} \varepsilon^i f_i^h \right) (\mathbf{v} \cdot \mathbf{n})^+ g \right\rangle d\mathbf{x}, \end{aligned}$$

or on a vu que

$$p_w(f_0^h + \varepsilon f_1^h + \dots + \varepsilon^n f_n^h) = p_w(f_0^h) + \varepsilon p_w(f_1^h) + \dots + \varepsilon^n p_w(f_n^h)$$

avec

$$\begin{cases} p_w(f_0^h) = \sigma_w(f_0^h) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w, \\ p_w(f_i^h) = \sigma_w(f_i^h) \quad \forall i \in \mathbb{N}^*, \\ \sigma_w(f_0^h) = -\frac{1}{2} \theta_w - \sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_0^h \right\rangle, \\ \sigma_w(f_i^h) = -\sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_i^h \right\rangle \quad \forall i \in \mathbb{N}^*, \end{cases}$$

et l'on a donc

$$\begin{aligned}
 & \sum_{K \in \mathbb{F}^\partial} \int_K \langle p_w(f^h) (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x} \\
 &= \sum_{K \in \mathbb{F}^\partial} \int_K \langle p_w(f_0^h) (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x} + \sum_{i \geq 1} \varepsilon^i \sum_{K \in \mathbb{F}^\partial} \int_K \langle p_w(f_i^h) (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x} \\
 &= \sum_{K \in \mathbb{F}^\partial} \int_K \langle p_w(f_0^h) (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x} + \sum_{i \geq 1} \varepsilon^i \sum_{K \in \mathbb{F}^\partial} \int_K \sigma_w(f_i^h) \langle (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x}.
 \end{aligned}$$

Cela donne finalement

➤ Termes en  $\varepsilon^{-1}$  : il n'y a pas de terme de bords, ce qui donne

$$\boxed{\sum_{C \in \mathbb{T}} \int_C \langle L(f_0^h) g M^{-1} \rangle d\mathbf{x} = 0 \quad \forall g \in \mathbb{E}_k.} \quad (3.22)$$

➤ Termes en  $\varepsilon^0$  :

$$\begin{aligned}
 \forall g \in \mathbb{E}_k, \quad & \sum_{C \in \mathbb{T}} \int_C \langle f_0^h \mathbf{v} \cdot \nabla_x g M^{-1} + L(f_1^h) g M^{-1} \rangle d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \langle \mathbf{v} \cdot \tilde{\mathbf{n}} [f_0^h] \llbracket g \rrbracket M^{-1} \rangle d\mathbf{x} \\
 &= - \sum_{K \in \mathbb{F}^\partial} \int_K \langle f_0^h g M^{-1} (\mathbf{v} \cdot \mathbf{n})^- \rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \langle p_w(f_0^h) (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x},
 \end{aligned} \quad (3.23)$$

ce qui peut s'écrire en utilisant la forme (3.21) de l'équation

$$\begin{aligned}
 \forall g \in \mathbb{E}_k, \quad & \sum_{C \in \mathbb{T}} \int_C \langle g \mathbf{v} \cdot \nabla_x f_0^h M^{-1} - L(f_1^h) g M^{-1} \rangle d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \langle \mathbf{v} \cdot \tilde{\mathbf{n}} [g] \llbracket f_0^h \rrbracket M^{-1} \rangle d\mathbf{x} \\
 &= - \sum_{K \in \mathbb{F}^\partial} \int_K \langle f_0^h g M^{-1} (\mathbf{v} \cdot \mathbf{n})^+ \rangle d\mathbf{x} + \sum_{K \in \mathbb{F}^\partial} \int_K \langle p_w(f_0^h) (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x}.
 \end{aligned} \quad (3.24)$$

➤ Termes d'ordres supérieurs :

$$\begin{aligned}
 & \forall i \in \mathbb{N}^*, \\
 \forall g \in \mathbb{E}_k, \quad & \sum_{C \in \mathbb{T}} \int_C \langle f_i^h \mathbf{v} \cdot \nabla_x g M^{-1} + L(f_{i+1}^h) g M^{-1} \rangle d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \langle \mathbf{v} \cdot \tilde{\mathbf{n}} [f_i^h] \llbracket g \rrbracket M^{-1} \rangle d\mathbf{x} \\
 &= - \sum_{K \in \mathbb{F}^\partial} \int_K \langle f_i^h g M^{-1} (\mathbf{v} \cdot \mathbf{n})^- \rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \sigma_w(f_i^h) \langle (\mathbf{v} \cdot \mathbf{n})^+ g \rangle d\mathbf{x},
 \end{aligned} \quad (3.25)$$

que l'on peut également écrire sous la forme

$$\begin{aligned}
 & \forall i \in \mathbb{N}^*, \\
 & \forall g \in \mathbb{E}_k, \quad \sum_{C \in \mathbb{T}} \int_C \left\langle g \mathbf{v} \cdot \nabla_x f_i^h M^{-1} - L(f_{i+1}^h) g M^{-1} \right\rangle d\mathbf{x} - \sum_{K \in \overset{\circ}{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [g] \llbracket f_i^h \rrbracket M^{-1} \right\rangle d\mathbf{x} \\
 & \quad = - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle f_i^h g M^{-1} (\mathbf{v} \cdot \mathbf{n})^+ \right\rangle d\mathbf{x} + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle \sigma_w(f_i^h) (\mathbf{v} \cdot \mathbf{n})^+ g \right\rangle d\mathbf{x}
 \end{aligned} \tag{3.26}$$

On analyse dans les sections suivantes les équations obtenues à chaque ordre.

### 3.4.2 Termes d'ordre -1

On considère l'équation (3.22) :  $\sum_{C \in \mathbb{T}} \int_C \left\langle L(f_0^h) g M^{-1} \right\rangle d\mathbf{x} = 0 \quad \forall g \in \mathbb{E}_k$ .

La fonction  $f^h$  étant un élément de  $\mathbb{E}_k$ , on considère que chacune de ses composantes l'est aussi :  $\forall i \in \mathbb{N}$ ,  $f_i^h \in \mathbb{E}_k$ . L'opérateur BGK linéarisé étant tel que  $\text{Im}(\mathbb{E}_k) \subset \mathbb{E}_k$ , on a aussi :

$$L(f_i^h) = \left[ \rho_i^h + \mathbf{u}_i^h \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T_i^h \right] M[1, \mathbf{0}, 1] - f_i^h \in \mathbb{E}_k, \quad \forall i \in \mathbb{N}$$

où les moments  $\rho_i^h$ ,  $\mathbf{u}_i^h$ ,  $T_i^h$  sont définis selon (3.2).

On peut donc choisir pour fonction-test  $g = L(f_0^h)$ , ce qui donne

$$\sum_{C \in \mathbb{T}} \int_C \left\langle L(f_0^h)^2 M^{-1} \right\rangle d\mathbf{x} = \|L(f_0^h)\|_{\mathbb{E}}^2 = 0,$$

soit  $L(f_0^h) = 0$ . La distribution  $f_0^h$  est donc une maxwellienne linéarisée :

$$f_0^h = \left[ \rho_0^h + \mathbf{u}_0^h \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T_0^h \right] M[1, \mathbf{0}, 1] := p_0^h M. \tag{3.27}$$

Les termes d'ordre suivant vont permettre de montrer la continuité et d'établir des conditions aux limites pour cette distribution.

### 3.4.3 Termes d'ordre 0

#### 3.4.3.1 Continuité de $f_0^h$ et conditions aux limites fortes

Dans cette partie, on cherche à établir au niveau discret les relations d'adhérence (3.3) vues au niveau continu. Sachant que  $f_0^h$  s'écrit  $f_0^h = p_0^h M$  avec  $p_0^h$  donné par (3.27), il suffit donc de montrer que  $p_0^h = p_w(f_0^h)$ . Pour cela, nous allons démontrer la propriété suivante :

**Proposition 4.**  $p_0^h$  est continue à travers les arêtes du maillage jusqu'au bord  $\mathbb{F}^\partial$ .

*Démonstration 4.* Considérons le terme  $p_w(f_0^h) = \sigma_w(f_0^h) + \mathbf{u}_w \cdot \mathbf{v} + \left(\frac{|\mathbf{v}|^2}{2} - \frac{3}{2}\right) \theta_w$ . Ce terme se compose de deux composantes de natures différentes. Si on le réécrit

$$\begin{aligned} p_w(f_0^h) &= \sigma_w(f_0^h) + \mathbf{u}_w \cdot \mathbf{v} + \left(\frac{|\mathbf{v}|^2}{2} - \frac{3}{2}\right) \theta_w \\ &= -\frac{1}{2} \theta_w - \underbrace{\sqrt{2\pi} \langle (\mathbf{v} \cdot \mathbf{n})^- f_0^h \rangle}_{:= \tilde{\sigma}_w(f_0^h)} + \mathbf{u}_w \cdot \mathbf{v} + \left(\frac{|\mathbf{v}|^2}{2} - \frac{3}{2}\right) \theta_w = \tilde{\sigma}_w(f_0^h) + \mathbf{u}_w \cdot \mathbf{v} + \left(\frac{|\mathbf{v}|^2}{2} - 2\right) \theta_w, \end{aligned}$$

le coefficient  $\tilde{\sigma}_w(f_0^h)$ , qui correspond à la masse de la maxwellienne linéarisée, est issu de la condition de flux de masse nul à travers la paroi et dépend donc de la solution du problème  $f_0^h$ .

En revanche, la composante  $\mathbf{u}_w \cdot \mathbf{v} + \left(\frac{|\mathbf{v}|^2}{2} - 2\right) \theta_w$  caractérise la perturbation des distributions induite par la paroi, et ne dépend que de la vitesse et de la température de cette paroi (ce sont des termes analogues à une condition de Dirichlet). L'étude de l'équation (3.23) va donc naturellement se traiter en considérant séparément ces deux composantes de la condition aux limites. On définit alors  $f_0^h$  par

$$f_0^h = \hat{f}_0^h + \check{f}_0^h,$$

où  $\hat{f}_0^h$  est définie par (3.23) (ou de façon équivalente par (3.24)) avec  $\mathbf{u}_w = \mathbf{0}$ ,  $\theta_w = 0$  (problème sans perturbation), et  $\check{f}_0^h$  est définie par les mêmes équations avec  $\tilde{\sigma}_w(\check{f}_0^h) = 0$  (problème sans masse). Nous montrons la continuité de ces deux composantes, puis nous concluons sur le problème complet par linéarité.

• **Problème sans perturbation** : on suppose dans un premier temps  $\mathbf{u}_w = \mathbf{0}$ ,  $\theta_w = 0$ .

Conformément à la relation (3.27), la solution  $\hat{f}_0^h$  du problème peut s'écrire  $\hat{f}_0^h = \hat{p}_0^h M$ . Pour montrer que  $\hat{p}_0^h$  est continue, nous allons montrer que son saut  $[[\hat{p}_0^h]]$  à travers les arêtes du maillages est nul. En choisissant la fonction-test  $g = \hat{p}_0^h M$  dans l'équation (3.23), on obtient

$$\begin{aligned} &\sum_{C \in \mathbb{T}} \int_C \left\langle \hat{f}_0^h \mathbf{v} \cdot \nabla_x \hat{p}_0^h + L(\hat{f}_1^h) \hat{p}_0^h \right\rangle d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [[\hat{p}_0^h]] [[\hat{f}_0^h]] \right\rangle d\mathbf{x} \\ &= - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \tilde{\sigma}_w(\hat{f}_0^h) \hat{p}_0^h M \right\rangle d\mathbf{x}. \end{aligned}$$

Les relations de conservation (3.16) et le fait que  $\hat{p}_0^h$  soit combinaison linéaire des invariants collisionnels  $1, \mathbf{v}, |\mathbf{v}|^2$  donnent

$$\sum_{C \in \mathbb{T}} \int_C \left\langle L(\hat{f}_1^h) \hat{p}_0^h \right\rangle d\mathbf{x} = 0 ; \quad (3.28)$$

en développant les deux autres termes du premier membre, il vient

$$\sum_{C \in \mathbb{T}} \int_C \left\langle \hat{f}_0^h \mathbf{v} \cdot \nabla_x \hat{p}_0^h \right\rangle d\mathbf{x} = \frac{1}{2} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [[(\hat{p}_0^h)^2]] M \right\rangle d\mathbf{x} - \frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle \mathbf{v} \cdot \mathbf{n} (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x}, \quad (3.29)$$

et

$$\sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [\hat{p}_0^h] [\hat{f}_0^h] \right\rangle d\mathbf{x} = \frac{1}{2} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [(\hat{p}_0^h)^2] M \right\rangle d\mathbf{x} + \frac{1}{2} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v} \cdot \tilde{\mathbf{n}}| [(\hat{p}_0^h)^2] M \right\rangle d\mathbf{x}, \quad (3.30)$$

ce qui donne finalement

$$\begin{aligned} \frac{1}{2} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v} \cdot \tilde{\mathbf{n}}| [(\hat{p}_0^h)^2] M \right\rangle d\mathbf{x} &= -\frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle \mathbf{v} \cdot \mathbf{n} (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x} + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x} \\ &\quad + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \tilde{\sigma}_w(\hat{f}_0^h) \hat{p}_0^h M \right\rangle d\mathbf{x} \\ &= -\frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x} + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \tilde{\sigma}_w(\hat{f}_0^h) \hat{p}_0^h M \right\rangle d\mathbf{x}. \end{aligned}$$

en remarquant que

$$-\frac{1}{2} \mathbf{v} \cdot \mathbf{n} + (\mathbf{v} \cdot \mathbf{n})^- = -\frac{1}{2} |\mathbf{v} \cdot \mathbf{n}|. \quad (3.31)$$

Dans la suite, nous allons montrer l'égalité entre  $(\hat{p}_0^h)$  et  $p_w(\hat{f}_0^h) = \tilde{\sigma}_w(\hat{f}_0^h)$  sur le bord. Pour cela, nous faisons apparaître le terme  $(\tilde{\sigma}_w(\hat{f}_0^h) - \hat{p}_0^h)^2$  dans le second terme du membre de droite de cette égalité en l'écrivant

$$\left\langle (\mathbf{v} \cdot \mathbf{n})^+ \tilde{\sigma}_w(\hat{f}_0^h) \hat{p}_0^h M \right\rangle = -\frac{1}{2} \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \left[ (\tilde{\sigma}_w(\hat{f}_0^h) - \hat{p}_0^h)^2 - (\tilde{\sigma}_w(\hat{f}_0^h))^2 - (\hat{p}_0^h)^2 \right] M \right\rangle.$$

Nous obtenons donc la relation :

$$\begin{aligned} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v} \cdot \tilde{\mathbf{n}}| [(\hat{p}_0^h)^2] M \right\rangle d\mathbf{x} &= - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x} \\ &\quad - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\tilde{\sigma}_w(\hat{f}_0^h) - \hat{p}_0^h)^2 M \right\rangle d\mathbf{x} + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x} \quad (3.32) \\ &\quad + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\tilde{\sigma}_w(\hat{f}_0^h))^2 M \right\rangle d\mathbf{x}. \end{aligned}$$

Nous allons maintenant montrer par une majoration du quatrième terme du membre de droite de l'égalité précédente que seul subsiste le second terme de ce membre, que nous avons fait apparaître, les autres se compensent. On écrit pour cela le coefficient  $\tilde{\sigma}_w(\hat{f}_0^h)$  sous la forme

$$\tilde{\sigma}_w(\hat{f}_0^h) = -\sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- \hat{f}_0^h \right\rangle = \int_{\mathbf{w} \cdot \mathbf{n} < 0} K(\mathbf{w}) \hat{p}_0^h(\mathbf{w}) M(\mathbf{w}) d\mathbf{w}, \quad \text{avec } K(\mathbf{w}) = -\sqrt{2\pi} \mathbf{w} \cdot \mathbf{n},$$

$$\text{et on obtient } \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\tilde{\sigma}_w(\hat{f}_0^h))^2 M \right\rangle = \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \left( \int_{\mathbf{w} \cdot \mathbf{n} < 0} K(\mathbf{w}) M(\mathbf{w}) \hat{p}_0^h(\mathbf{w}) d\mathbf{w} \right)^2 M(\mathbf{v}) \right\rangle.$$

$$\text{Or } \int_{\mathbf{w} \cdot \mathbf{n} < 0} K(\mathbf{w}) M(\mathbf{w}) d\mathbf{w} = -\sqrt{2\pi} \int_{\mathbf{w} \cdot \mathbf{n} < 0} \mathbf{w} \cdot \mathbf{n} M(\mathbf{w}) d\mathbf{w} = 1,$$

et on peut utiliser l'inégalité de Jensen [41], ce qui donne

$$\begin{aligned}
\left\langle (\mathbf{v} \cdot \mathbf{n})^+ \left( \tilde{\sigma}_w(\hat{f}_0^h) \right)^2 M \right\rangle &\leq \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \int_{\mathbf{w} \cdot \mathbf{n} < 0} K(\mathbf{w}) M(\mathbf{w}) (\hat{p}_0^h(\mathbf{w}))^2 d\mathbf{w} M(\mathbf{v}) \right\rangle \\
&= \int_{\mathbf{w} \cdot \mathbf{n} < 0} \left\langle (\mathbf{v} \cdot \mathbf{n})^+ K(\mathbf{w}) (\mathbf{w} \cdot \mathbf{n})^{-1} M(\mathbf{v}) \right\rangle \mathbf{w} \cdot \mathbf{n} (\hat{p}_0^h(\mathbf{w}))^2 M(\mathbf{w}) d\mathbf{w} \\
&= -\sqrt{2\pi} \int_{\mathbf{w} \cdot \mathbf{n} < 0} \left\langle (\mathbf{v} \cdot \mathbf{n})^+ M(\mathbf{v}) \right\rangle \mathbf{w} \cdot \mathbf{n} (\hat{p}_0^h(\mathbf{w}))^2 M(\mathbf{w}) d\mathbf{w} \\
&= -\int_{\mathbf{w} \cdot \mathbf{n} < 0} \mathbf{w} \cdot \mathbf{n} (\hat{p}_0^h(\mathbf{w}))^2 M(\mathbf{w}) d\mathbf{w} \\
&= -\left\langle (\mathbf{v} \cdot \mathbf{n})^- (\hat{p}_0^h)^2 M \right\rangle.
\end{aligned}$$

En appliquant cette inégalité à la relation (4.13), on obtient

$$\begin{aligned}
\sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v} \cdot \tilde{\mathbf{n}}| [\hat{p}_0^h]^2 M \right\rangle d\mathbf{x} &\leq - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x} \\
&\quad - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \left( \tilde{\sigma}_w(\hat{f}_0^h) - \hat{p}_0^h \right)^2 M \right\rangle d\mathbf{x} \\
&\quad + \underbrace{\sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x}}_{= \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| (\hat{p}_0^h)^2 M \right\rangle d\mathbf{x}} \\
&= - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \left( \tilde{\sigma}_w(\hat{f}_0^h) - \hat{p}_0^h \right)^2 M \right\rangle d\mathbf{x} \leq 0,
\end{aligned}$$

$$\text{soit } 0 \leq \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v} \cdot \tilde{\mathbf{n}}| [\hat{p}_0^h]^2 M \right\rangle d\mathbf{x} \leq - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \left( \tilde{\sigma}_w(\hat{f}_0^h) - \hat{p}_0^h \right)^2 M \right\rangle d\mathbf{x} \leq 0.$$

Les deux termes de cette inégalité sont donc nuls ; le premier donne  $[\hat{p}_0^h] = 0$ , et donc  $\hat{f}_0^h$  est continue. Le second terme donne  $\hat{p}_0^h = \tilde{\sigma}_w(\hat{f}_0^h)$  sur le bord du domaine.  $\square$

• **Problème sans masse** : on considère maintenant le cas  $\tilde{\sigma}_w = 0$ ,  $\mathbf{u}_w \neq \mathbf{0}$ ,  $\theta_w \neq 0$ .

Comme dans le cas précédent, on a  $\check{f}_0^h = \check{p}_0^h M$ .

La condition aux limites  $\alpha(\mathbf{v}) M = \left[ \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - 2 \right) \theta_w \right] M$  étant continue, on peut construire un relèvement continu  $l_0^h M$  de cette condition. On choisit alors dans l'équation (3.23) la fonction-test  $g = (\check{p}_0^h - l_0^h) M$ , ce qui donne

$$\begin{aligned}
\sum_{C \in \mathbb{T}} \int_C \left\langle \check{f}_0^h \mathbf{v} \cdot \nabla_x \check{p}_0^h + L(f_1^h) \check{p}_0^h \right\rangle d\mathbf{x} &- \sum_{C \in \mathbb{T}} \int_C \left\langle \check{f}_0^h \mathbf{v} \cdot \nabla_x l_0^h + L(f_1^h) l_0^h \right\rangle d\mathbf{x} \\
&- \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [\check{f}_0^h] [\check{p}_0^h] \right\rangle d\mathbf{x} + \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [\check{f}_0^h] \underbrace{[l_0^h]}_{=0} \right\rangle d\mathbf{x}
\end{aligned}$$

$$\begin{aligned}
&= - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- (\check{p}_0^h)^2 M \right\rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \alpha(\mathbf{v}) \check{p}_0^h M \right\rangle d\mathbf{x} \\
&\quad + \underbrace{\sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- \check{f}_0^h l_0^h \right\rangle d\mathbf{x}}_{\text{}} + \underbrace{\sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \alpha(\mathbf{v}) M l_0^h \right\rangle d\mathbf{x}}_{\text{}} . \\
&= \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- \check{f}_0^h \alpha(\mathbf{v}) \right\rangle d\mathbf{x} = \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x}
\end{aligned}$$

On suppose que  $l_0^h$  est de la même forme que  $\alpha(\mathbf{v})$ , soit

$$l_0^h = \mathbf{u} \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - 2 \right) T,$$

ce qui donne  $\sum_{C \in \mathbb{T}} \int_C \left\langle L(f_1^h) l_0^h \right\rangle d\mathbf{x} = 0$ . Avec les mêmes transformations que dans le problème sans perturbation (voir les relations (3.29) et (4.14)), on obtient

$$\begin{aligned}
&- \sum_{C \in \mathbb{T}} \int_C \left\langle \check{f}_0^h \mathbf{v} \cdot \nabla_x l_0^h \right\rangle d\mathbf{x} - \frac{1}{2} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v} \cdot \tilde{\mathbf{n}}| [\check{p}_0^h]^2 M \right\rangle d\mathbf{x} - \frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle \mathbf{v} \cdot \mathbf{n} (\check{p}_0^h)^2 M \right\rangle d\mathbf{x} \\
&= - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- (\check{p}_0^h)^2 M \right\rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ \alpha(\mathbf{v}) \check{p}_0^h M \right\rangle d\mathbf{x} \\
&\quad + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- \check{f}_0^h \alpha(\mathbf{v}) \right\rangle d\mathbf{x} + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x}.
\end{aligned}$$

Pour simplifier cette expression, on combine le premier terme du membre de gauche et le dernier du membre de droite en utilisant de nouveau la relation (3.31), ainsi que les second et troisième terme du membre de droite en utilisant la relation  $(\mathbf{v} \cdot \mathbf{n})^- - (\mathbf{v} \cdot \mathbf{n})^+ = -|\mathbf{v} \cdot \mathbf{n}|$ , pour trouver

$$\begin{aligned}
&- \sum_{C \in \mathbb{T}} \int_C \left\langle \check{f}_0^h \mathbf{v} \cdot \nabla_x l_0^h \right\rangle d\mathbf{x} - \frac{1}{2} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v} \cdot \tilde{\mathbf{n}}| [\check{p}_0^h]^2 M \right\rangle d\mathbf{x} \\
&= \frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| (\check{p}_0^h)^2 M \right\rangle d\mathbf{x} - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| \alpha(\mathbf{v}) \check{p}_0^h M \right\rangle d\mathbf{x} \\
&\quad + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x} \\
&= \frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| (\check{p}_0^h - \alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x} - \frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| (\alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x} \\
&\quad + \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^+ (\alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x} \\
&= \frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| (\check{p}_0^h - \alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x} + \frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle \mathbf{v} \cdot \mathbf{n} (\alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x}.
\end{aligned}$$

Enfin, on remarque que le dernier terme de l'expression précédente est nul :

$$\sum_{K \in \mathbb{F}^\partial} \int_K \left\langle \mathbf{v} \cdot \mathbf{n} (\alpha(\mathbf{v}))^2 M \right\rangle d\mathbf{x} = \sum_{K \in \mathbb{F}^\partial} \int_K \theta_w \underbrace{\mathbf{u}_w \cdot \mathbf{n}}_{=0} d\mathbf{x} = 0,$$



et on a donc finalement

$$\begin{aligned} \frac{1}{2} \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v} \cdot \tilde{\mathbf{n}}| [\tilde{p}_0^h]^2 M \right\rangle d\mathbf{x} + \frac{1}{2} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle |\mathbf{v} \cdot \mathbf{n}| \left( \tilde{p}_0^h - \alpha(\mathbf{v}) \right)^2 M \right\rangle d\mathbf{x} \\ = - \sum_{C \in \mathbb{T}} \int_C \left\langle \check{f}_0^h \mathbf{v} \cdot \nabla_x l_0^h \right\rangle d\mathbf{x}. \end{aligned} \quad (3.33)$$

Nous souhaitons montrer la continuité de  $\tilde{p}_0^h$ , c'est-à-dire  $[\tilde{p}_0^h] = 0$ , ainsi que l'adhérence à la paroi, soit  $\check{p}_{0|\Gamma}^h = \alpha(\mathbf{v})$ . Ces deux résultats sont obtenus si les deux membres de la relation (3.33) sont nuls, mais le membre de droite n'a pas de signe *a priori*. À ce stade, une hypothèse est donc nécessaire. On se place dans le cas particulier évoqué au paragraphe 3.2.3 : la paroi est supposée indéformable, et donc la vitesse  $\mathbf{u}_w$ , purement tangentielle par hypothèse, est uniforme. On va également devoir supposer  $\theta_w$  uniforme. On peut alors choisir  $l_0^h \equiv \alpha(\mathbf{v})$ , et donc  $\nabla_x l_0^h \equiv \mathbf{0}$ , et ainsi

$$\sum_{C \in \mathbb{T}} \int_C \left\langle \check{f}_0^h \mathbf{v} \cdot \nabla_x l_0^h \right\rangle d\mathbf{x} = 0.$$

On a ainsi le résultat voulu :  $\tilde{p}_0^h$  (et donc  $\check{f}_0^h$ ) est continue et

$$\check{f}_0^h = \alpha(\mathbf{v}) M = \left[ \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - 2 \right) \theta_w \right] M$$

à la paroi.  $\square$

☞ NB : on verra dans la section suivante que si l'hypothèse supposant  $\mathbf{u}_w$  uniforme n'est pas toujours indispensable, on ne peut pas se passer de la même hypothèse sur  $\theta_w$ .  $\square$

• **Problème complet** : par linéarité du problème complet, la solution  $f_0^h$  de ce problème peut s'écrire comme la somme de la solution  $\hat{f}_0^h$  du problème sans perturbation et de la solution  $\check{f}_0^h$  du problème sans masse, ce que nous démontrons ci-dessous.

Tout d'abord,  $\hat{f}_0^h$  et  $\check{f}_0^h$  étant continues,  $f_0^h$  est continue.

Par ailleurs, on obtient à la paroi :

$$p_0^h = \hat{p}_0^h + \check{p}_0^h = \tilde{\sigma}_w(\hat{f}_0^h) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - 2 \right) \theta_w,$$

et on souhaite obtenir

$$p_0^h = \tilde{\sigma}_w(f_0^h) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - 2 \right) \theta_w.$$

Il suffit donc de montrer que

$$\tilde{\sigma}_w(\check{f}_0^h) = -\sqrt{2\pi} \left\langle (\mathbf{v} \cdot \mathbf{n})^- \left[ \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - 2 \right) \theta_w \right] M \right\rangle = 0.$$

En tenant compte du fait que  $\mathbf{u}_w$  est tangente à la paroi, ce résultat est obtenu par un calcul direct (voir Annexe D).

Finalement, on a

$$p_0^h = \tilde{\sigma}_w(\hat{f}_0^h) + \tilde{\sigma}_w(\check{f}_0^h) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - 2 \right) \theta_w = \sigma_w(f_0^h) + \mathbf{u}_w \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) \theta_w,$$

donc  $f_0^h = \hat{f}_0^h + \check{f}_0^h$  vérifie bien l'équation (3.23) et l'on a les conditions d'adhérence fortes :

$$\boxed{\begin{cases} \rho_0^h = \sigma_w(f_0^h), \\ \mathbf{u}_0^h = \mathbf{u}_w, \\ T_0^h = \theta_w. \end{cases}} \quad \square \quad (3.34)$$

### 3.4.3.2 Remarque sur l'hypothèse d'uniformité de $\mathbf{u}_w$ et $\theta_w$

Nous allons montrer en considérant quelques cas simple que l'hypothèse d'uniformité faite dans la démonstration précédente sur les perturbations à la paroi  $\mathbf{u}_w$  et  $\theta_w$  se justifient.

- **Écoulement de Couette**

On considère un écoulement de gaz raréfié entre deux plaques planes animées de vitesses différentes, comme montré sur la figure 3.3.

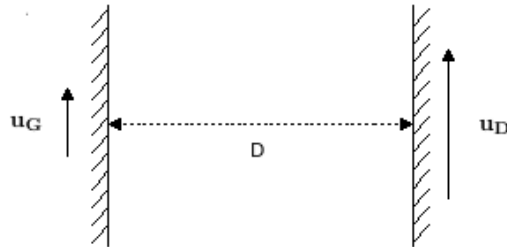


FIGURE 3.3: Écoulement de Couette

On a  $\theta_w = 0$ , soit  $\alpha(\mathbf{v}) = \mathbf{u}_w \cdot \mathbf{v}$ , avec  $\mathbf{u}_w = \mathbf{u}_G$  sur la plaque gauche, et  $\mathbf{u}_w = \mathbf{u}_D$  sur la plaque droite. On choisit alors un relèvement de la forme

$$l_0^h = \mathbf{u} \cdot \mathbf{v}$$

avec  $\mathbf{u} = u(x)\mathbf{e}_y$  une fonction variant continument de  $\mathbf{u}_G$  à  $\mathbf{u}_D$ . On obtient alors très simplement  $\mathbf{v} \cdot \nabla_x l_0^h = u'(x)v_x v_y$ , et

$$\left\langle \check{f}_0^h \mathbf{v} \cdot \nabla_x l_0^h \right\rangle = 0 \quad \text{par imparité de tous les termes.}$$

Ce résultat montre que pour une configuration présentant plusieurs parois, on obtient le résultat voulu même si les différentes parois possèdent des vitesses différentes.

### • Écoulement de Fourier

L'écoulement présente la même configuration que dans le cas de Couette, mais cette fois les plaques sont immobiles, et leur température est différente : la plaque gauche est à la température  $T_G$ , la plaque droite à la température  $T_D$ . On a donc  $\alpha(\mathbf{v}) = \left(\frac{|\mathbf{v}|^2}{2} - 2\right)\theta_w$ , et on pose

$$l_0^h = \left(\frac{|\mathbf{v}|^2}{2} - 2\right)T,$$

avec  $T = T(x)$  variant continument de  $T_G$  à  $T_D$ . On obtient ainsi  $\mathbf{v} \cdot \nabla_x l_0^h = T'(x)v_x \left(\frac{|\mathbf{v}|^2}{2} - 2\right)$ , et

$$\left\langle \check{f}_0^h \mathbf{v} \cdot \nabla_x l_0^h \right\rangle = \left\langle v_x^2 \left(\frac{|\mathbf{v}|^2}{2} - 2\right) M \right\rangle \check{\mathbf{u}}_{0x}^h T'(x) = \frac{1}{2} \check{\mathbf{u}}_{0x}^h T'(x).$$

Le terme  $\check{\mathbf{u}}_{0x}^h$  n'ayant pas de raison d'être nul, on doit alors supposer  $T_G = T_D$  de façon à pouvoir imposer  $T = \text{cste}$ , et ainsi  $T'(x) = 0$  pour obtenir le résultat souhaité.

L'analyse du glissement thermique donne le même résultat. Ces exemples montrent bien que l'hypothèse  $\theta_w = \text{cste}$  semble bien nécessaire. Elle est certes très restrictive, mais nous ne voyons pas pour le moment comment nous en passer.

### 3.4.3.3 Équations de continuité et de Boussinesq

Les conditions d'adhérence (3.34) et la continuité de  $f_0^h$  entraînent l'annulation des termes d'arêtes dans la forme (3.24) de l'équation, qui s'écrit alors

$$\sum_{C \in \mathbb{T}} \int_C \left\langle \left( \mathbf{v} \cdot \nabla_x f_0^h - L(f_1^h) \right) g M^{-1} \right\rangle d\mathbf{x} = 0. \quad (3.35)$$

A partir de cette forme très simple, on peut déduire les formes faibles de l'équation de continuité et de la relation de Boussinesq. On va utiliser l'équation (3.35) avec deux types de fonction-test :

➤ On choisit d'abord une fonction-test de la forme  $g = \varphi M$ , avec  $\varphi \in \mathbb{X}_k$  (soit  $\varphi$  indépendante de  $\mathbf{v}$ ). On a alors

$$\begin{cases} \sum_{C \in \mathbb{T}} \int_C \left\langle g \mathbf{v} \cdot \nabla_x f_0^h M^{-1} \right\rangle d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \left\langle \mathbf{v} \cdot \nabla_x f_0^h \right\rangle \varphi d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \nabla \cdot \mathbf{u}_0^h \varphi d\mathbf{x}, \\ \sum_{C \in \mathbb{T}} \int_C \left\langle L(f_1^h) g M^{-1} \right\rangle d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle L(f_1^h) \right\rangle}_{=0} \varphi d\mathbf{x} = 0. \end{cases}$$

On obtient donc la forme faible de l'équation de continuité

$$\boxed{\sum_{C \in \mathbb{T}} \int_C \nabla \cdot \mathbf{u}_0^h \varphi d\mathbf{x} = 0 \quad \forall \varphi \in \mathbb{X}_k.} \quad (3.36)$$

➤ On prend ensuite  $g = \mathbf{v} \cdot \boldsymbol{\varphi} M$ ,  $\boldsymbol{\varphi} \in \mathbb{Y}_k$ , ce qui donne

$$\begin{cases} \sum_{C \in \mathbb{T}} \int_C \left\langle g \mathbf{v} \cdot \nabla_x f_0^h M^{-1} \right\rangle d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \left\langle \mathbf{v} (\mathbf{v} \cdot \nabla_x f_0^h) \right\rangle \cdot \boldsymbol{\varphi} d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \nabla (\rho_0^h + T_0^h) \cdot \boldsymbol{\varphi} d\mathbf{x} = 0, \\ \sum_{C \in \mathbb{T}} \int_C \left\langle L(f_1^h) g M^{-1} \right\rangle d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle \mathbf{v} L(f_1^h) \right\rangle}_{=0} \cdot \boldsymbol{\varphi} d\mathbf{x} = 0, \end{cases}$$

soit la forme faible de la relation de Boussinesq

$$\boxed{\sum_{C \in \mathbb{T}} \int_C \nabla (\rho_0^h + T_0^h) \cdot \boldsymbol{\varphi} d\mathbf{x} = 0 \quad \forall \boldsymbol{\varphi} \in \mathbb{Y}_k.} \quad (3.37)$$

➤ Comme dans le cas continu, le choix de fonction-test  $g = \varphi |\mathbf{v}|^2 M$  donne le même résultat que  $g = \varphi M$ ,  $\varphi \in \mathbb{X}_k$ .

Enfin, en utilisant  $L(f_1^h) = \Pi(f_1^h) - f_1^h$  dans (3.35), on exprime la distribution  $f_1^h$  sous forme faible :

$$\boxed{\sum_{C \in \mathbb{T}} \int_C \left\langle f_1^h g M^{-1} \right\rangle d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \left\langle -\mathbf{v} \cdot \nabla_x (f_0^h) g M^{-1} + \Pi(f_1^h) g M^{-1} \right\rangle d\mathbf{x} \quad \forall g \in \mathbb{E}_k.} \quad (3.38)$$

On passe ensuite, comme pour les distributions continues, aux termes d'ordre 1 en  $\varepsilon$  pour établir la forme faible des équations de Stokes et de Laplace.

### 3.4.4 Termes d'ordre 1

A la différence de ce qui se passe pour  $f_0^h$ , nous n'avons pas d'expression forte pour  $f_1^h$ . Il est donc impossible de procéder comme dans le cas précédent pour établir la continuité de  $f_1^h$  ou des conditions aux limites fortes. Le terme sur les arêtes intérieures de l'équation (3.26) est donc a priori non nul ; on choisit donc d'utiliser l'équation (3.25) en imposant la continuité des fonctions-test pour faire disparaître ce terme. On procède alors comme à l'ordre 0, en choisissant pour fonctions-test des monômes en fonction de  $\mathbf{v}$ .

L'équation (3.25) s'écrit à l'ordre 1

$$\begin{aligned} \forall g \in \mathbb{E}_k, \quad & \sum_{C \in \mathbb{T}} \int_C \left\langle f_1^h \mathbf{v} \cdot \nabla_x g M^{-1} + L(f_2^h) g M^{-1} \right\rangle d\mathbf{x} - \sum_{K \in \overset{\circ}{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} [f_1^h] [g] M^{-1} \right\rangle d\mathbf{x} \\ & = - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle f_1^h g M^{-1} (\mathbf{v} \cdot \mathbf{n})^- \right\rangle d\mathbf{x} + \sqrt{2\pi} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_1^h \right\rangle \left\langle (\mathbf{v} \cdot \mathbf{n})^+ g \right\rangle d\mathbf{x}. \end{aligned} \quad (3.39)$$

➤ Pour obtenir l'équation de continuité sur  $\mathbf{u}_1^h$ , on choisit comme à l'ordre précédent  $g = \varphi M$ ,  $\varphi \in \mathbb{X}_k \cap \mathcal{C}^0(\Omega)$  : le premier membre donne

$$\begin{aligned} & \sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle f_1^h \mathbf{v} \cdot \nabla_x \varphi \right\rangle}_{= \left\langle \mathbf{v} \cdot f_1^h \right\rangle \nabla_x \varphi} d\mathbf{x} + \sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle L(f_2^h) \right\rangle}_{= 0} \varphi d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}}[f_1^h] \underbrace{\llbracket \varphi \rrbracket}_{= 0} \right\rangle d\mathbf{x}, \\ & = \left\langle \mathbf{v} \cdot f_1^h \right\rangle \nabla_x \varphi \end{aligned}$$

et le second

$$- \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle f_1^h (\mathbf{v} \cdot \mathbf{n})^- \right\rangle \varphi d\mathbf{x} + \sqrt{2\pi} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_1^h \right\rangle \underbrace{\left\langle M(\mathbf{v} \cdot \mathbf{n})^+ \right\rangle}_{= \frac{1}{\sqrt{2\pi}}} \varphi d\mathbf{x} = 0.$$

On obtient donc la forme faible de l'équation de continuité

$$\boxed{\sum_{C \in \mathbb{T}} \int_C \mathbf{u}_1^h \cdot \nabla_x \varphi d\mathbf{x} = 0 \quad \forall \varphi \in \mathbb{X}_k \cap \mathcal{C}^0(\Omega).} \quad (3.40)$$

➤ Pour obtenir l'équation de Stokes, on choisit ensuite naturellement une fonction-test de la forme  $g = \mathbf{v} \cdot \varphi M$ ,  $\varphi \in \mathbb{Y}_k \cap \mathcal{C}^0(\Omega)$ .

Les termes du second membre de (3.39) ne disparaissent pas naturellement comme pour l'équation de continuité. On se restreint alors à des fonctions-test nulles sur le bord du domaine, c'est-à-dire  $g = \mathbf{v} \cdot \varphi M$ ,  $\varphi \in \mathbb{Y}_k \cap \mathcal{C}_0^0(\Omega)$ , soit

$$\sum_{C \in \mathbb{T}} \int_C \left\langle \underbrace{f_1^h \mathbf{v} \cdot \nabla_x (\mathbf{v} \cdot \varphi)}_{= f_1^h (\mathbf{v} \otimes \mathbf{v}) : \nabla_x \varphi} \right\rangle d\mathbf{x} + \sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle \mathbf{v} L(f_2^h) \right\rangle}_{= 0} \cdot \varphi d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}}[f_1^h] \underbrace{\mathbf{v} \cdot \llbracket \varphi \rrbracket}_{= 0} \right\rangle d\mathbf{x} = 0.$$

On pose  $\bar{\bar{J}}_1 = \left\langle (\mathbf{v} \otimes \mathbf{v}) f_1^h \right\rangle$ , ce qui donne :

$$\sum_{C \in \mathbb{T}} \int_C \bar{\bar{J}}_1 : \nabla_x \varphi d\mathbf{x} = 0 \quad \forall \varphi \in \mathbb{Y}_k \cap \mathcal{C}_0^0(\Omega).$$

On cherche ensuite à exprimer  $\bar{\bar{J}}_1$  sous forme faible, ce qui va nous permettre d'aboutir à une formulation mixte de l'équation de Stokes. L'équation indiquée pour cela est la forme faible de  $f_1^h$  (3.38), avec la fonction-test  $g = (\mathbf{v} \otimes \mathbf{v}) : \bar{\bar{\psi}} M$ ,  $\bar{\bar{\psi}} \in \mathbb{Y}_k \otimes \mathbb{Y}_k$  :

$$\sum_{C \in \mathbb{T}} \int_C \left\langle f_1^h (\mathbf{v} \otimes \mathbf{v}) \right\rangle : \bar{\bar{\psi}} d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \left\langle -\mathbf{v} \cdot \nabla_x (f_0^h) (\mathbf{v} \otimes \mathbf{v}) + \Pi(f_1^h) (\mathbf{v} \otimes \mathbf{v}) \right\rangle : \bar{\bar{\psi}} d\mathbf{x},$$

or

$$\begin{cases} \left\langle \mathbf{v} \cdot \nabla_x (f_0^h) (\mathbf{v} \otimes \mathbf{v}) \right\rangle : \bar{\bar{\psi}} = \nabla \cdot \mathbf{u}_0^h \operatorname{Tr}(\bar{\bar{\psi}}) + \left[ \nabla_x \mathbf{u}_0^h + (\nabla_x \mathbf{u}_0^h)^T \right] : \bar{\bar{\psi}}, \\ \left\langle \Pi(f_1^h) (\mathbf{v} \otimes \mathbf{v}) \right\rangle : \bar{\bar{\psi}} = (\rho_1^h + T_1^h) \operatorname{Tr}(\bar{\bar{\psi}}) = p_1^h \operatorname{Tr}(\bar{\bar{\psi}}). \end{cases}$$

Avec l'équation d'incompressibilité (3.36), on obtient  $\sum_{C \in \mathbb{T}} \int_C \nabla \cdot \mathbf{u}_0^h \operatorname{Tr}(\bar{\bar{\psi}}) d\mathbf{x} = 0$ , ce qui donne

$$\sum_{C \in \mathbb{T}} \int_C \bar{\bar{J}}_1 : \bar{\bar{\psi}} d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \left[ p_1^h \text{Tr}(\bar{\bar{\psi}}) - \left( \nabla_x \mathbf{u}_0^h + \left( \nabla_x \mathbf{u}_0^h \right)^T \right) : \bar{\bar{\psi}} \right] d\mathbf{x},$$

et on obtient la forme mixte de l'équation de Stokes :

$$\left\{ \begin{array}{l} \sum_{C \in \mathbb{T}} \int_C \bar{\bar{J}}_1 : \nabla_x \boldsymbol{\varphi} d\mathbf{x} = 0 \quad \forall \boldsymbol{\varphi} \in \mathbb{Y}_k \cap \mathcal{C}_0^0(\Omega), \\ \sum_{C \in \mathbb{T}} \int_C \bar{\bar{J}}_1 : \bar{\bar{\psi}} d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \left[ p_1^h \text{Tr}(\bar{\bar{\psi}}) - \left[ \nabla_x \mathbf{u}_0^h + \left( \nabla_x \mathbf{u}_0^h \right)^T \right] : \bar{\bar{\psi}} \right] d\mathbf{x} \quad \forall \bar{\bar{\psi}} \in \mathbb{Y}_k \otimes \mathbb{Y}_k. \end{array} \right. \quad (3.41)$$

Cette équation peut se mettre sous forme non mixte. On pose  $\bar{\bar{\psi}} = \nabla_x \boldsymbol{\varphi}$  dans le système (3.41), avec  $\boldsymbol{\varphi} \in \mathbb{Y}_k \cap \mathcal{C}_0^0(\Omega)$ , et on obtient par égalité des deux premiers membres

$$\sum_{C \in \mathbb{T}} \int_C \left[ p_1^h \nabla \cdot \boldsymbol{\varphi} - \left( \nabla_x \mathbf{u}_0^h + \left( \nabla_x \mathbf{u}_0^h \right)^T \right) : \nabla_x \boldsymbol{\varphi} \right] d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \bar{\bar{J}}_1 : \nabla_x \boldsymbol{\varphi} d\mathbf{x} = 0 ;$$

$$\begin{aligned} \text{or } \sum_{C \in \mathbb{T}} \int_C \left( \nabla_x \mathbf{u}_0^h \right)^T : \nabla_x \boldsymbol{\varphi} d\mathbf{x} &= \sum_{C \in \mathbb{T}} \int_C \left( \frac{\partial \mathbf{u}_{0j}^h}{\partial x_i} \frac{\partial \varphi_i}{\partial x_j} \right) d\mathbf{x} \\ &= - \sum_{C \in \mathbb{T}} \int_C \left( \mathbf{u}_{0j}^h \frac{\partial^2 \varphi_i}{\partial x_i \partial x_j} \right) d\mathbf{x} + \sum_{C \in \mathbb{T}} \int_{\partial C} \mathbf{u}_{0j}^h \frac{\partial \varphi_i}{\partial x_j} n_i d\mathbf{x} \\ &= - \sum_{C \in \mathbb{T}} \int_C \mathbf{u}_0^h \nabla_x (\nabla \cdot \boldsymbol{\varphi}) d\mathbf{x} + \sum_{K \in \mathring{\mathbb{F}}} \int_K \llbracket \nabla_x \boldsymbol{\varphi} : (\tilde{\mathbf{n}} \otimes \mathbf{u}_0^h) \rrbracket d\mathbf{x}. \end{aligned}$$

En effectuant une intégration par parties de l'équation (3.36) avec  $\boldsymbol{\varphi} = \nabla \cdot \boldsymbol{\varphi}$  continue (c'est-à-dire  $\boldsymbol{\varphi}$  de classe  $\mathcal{C}^1$ ), on voit que le premier terme du second membre de l'égalité précédente est nul. En outre,  $\nabla_x \boldsymbol{\varphi}$  est alors continue et le second terme est également nul (la continuité de  $f_0^h$  entraîne celle de  $\mathbf{u}_0^h$ ). On aboutit à la forme non mixte

$$\left\{ \sum_{C \in \mathbb{T}} \int_C \left[ p_1^h \nabla \cdot \boldsymbol{\varphi} - \nabla_x \mathbf{u}_0^h : \nabla_x \boldsymbol{\varphi} \right] d\mathbf{x} = 0 \quad \forall \boldsymbol{\varphi} \in \mathbb{Y}_k \cap \mathcal{C}_0^1(\Omega). \right. \quad (3.42)$$

➤ Pour obtenir l'équation de Laplace on prend, toujours afin d'éliminer les termes de bord,  $g = \varphi |\mathbf{v}|^2 M$ ,  $\varphi \in \mathbb{X}_k \cap \mathcal{C}_0^0(\Omega)$  dans (3.39) pour trouver :

$$\sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle f_1^h \mathbf{v} \cdot \nabla_x (\varphi |\mathbf{v}|^2) \right\rangle}_{= \left\langle f_1^h \mathbf{v} |\mathbf{v}|^2 \right\rangle \cdot \nabla_x \varphi} d\mathbf{x} + \sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle |\mathbf{v}|^2 L(f_2^h) \right\rangle}_{= 0} \varphi d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle |\mathbf{v}|^2 \mathbf{v} \cdot \tilde{\mathbf{n}} [f_1^h] \underbrace{\llbracket \varphi \rrbracket}_{= 0} \right\rangle d\mathbf{x} = 0.$$

On pose  $\mathbf{J}'_1 = \left\langle \mathbf{v} |\mathbf{v}|^2 f_1^h \right\rangle$ , ce qui donne

$$\sum_{C \in \mathbb{T}} \int_C \mathbf{J}'_1 \cdot \nabla_x \varphi d\mathbf{x} = 0 \quad \forall \varphi \in \mathbb{X}_k \cap \mathcal{C}^0(\Omega).$$

Comme pour l'équation de Stokes, on va aboutir à une formulation mixte : en choisissant  $g = |\mathbf{v}|^2 \mathbf{v} \cdot \boldsymbol{\psi} M$ ,  $\boldsymbol{\psi} \in \mathbb{Y}_k$  dans la forme faible (3.38), on obtient

$$\sum_{C \in \mathbb{T}} \int_C \left\langle f_1^h |\mathbf{v}|^2 \mathbf{v} \right\rangle \cdot \boldsymbol{\psi} d\mathbf{x} = - \sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle |\mathbf{v}|^2 (\mathbf{v} \otimes \mathbf{v}) \nabla_x (f_0^h) \right\rangle}_{= 5 \nabla (\rho_0^h + 2T_0^h)} \cdot \boldsymbol{\psi} d\mathbf{x} + \sum_{C \in \mathbb{T}} \int_C \underbrace{\left\langle \Pi(f_1^h) \mathbf{v} |\mathbf{v}|^2 \right\rangle}_{= 5 \mathbf{u}_1^h} \cdot \boldsymbol{\psi} d\mathbf{x}.$$

On utilise la relation de Boussinesq (3.37) pour éliminer  $\rho_0^h$  de la relation précédente et obtenir :

$$\sum_{C \in \mathbb{T}} \int_C \mathbf{J}'_1 \cdot \boldsymbol{\psi} d\mathbf{x} = -5 \sum_{C \in \mathbb{T}} \int_C \nabla T_0^h \cdot \boldsymbol{\psi} d\mathbf{x} + 5 \sum_{C \in \mathbb{T}} \int_C \mathbf{u}_1^h \cdot \boldsymbol{\psi} d\mathbf{x}.$$

On obtient alors la formulation mixte de l'équation de Laplace

$$\left\{ \begin{array}{l} \sum_{C \in \mathbb{T}} \int_C \mathbf{J}'_1 \cdot \nabla_x \varphi d\mathbf{x} = 0 \quad \forall \varphi \in \mathbb{X}_k \cap \mathcal{C}_0^0(\Omega), \\ \sum_{C \in \mathbb{T}} \int_C \mathbf{J}'_1 \cdot \boldsymbol{\psi} d\mathbf{x} = 5 \sum_{C \in \mathbb{T}} \int_C [\mathbf{u}_1^h - \nabla T_0^h] \cdot \boldsymbol{\psi} d\mathbf{x} \quad \forall \boldsymbol{\psi} \in \mathbb{Y}_k. \end{array} \right.$$

On met cette équation sous forme non mixte de la même façon que précédemment : en choisissant  $\boldsymbol{\psi} = \nabla_x \varphi$ ,  $\varphi \in \mathbb{X}_k \cap \mathcal{C}_0^0(\Omega)$  dans le système obtenu, on a

$$\underbrace{5 \sum_{C \in \mathbb{T}} \int_C \mathbf{u}_1^h \cdot \nabla_x \varphi d\mathbf{x}}_{= 0 \text{ selon (3.40)}} - 5 \sum_{C \in \mathbb{T}} \int_C \nabla T_0^h \cdot \nabla_x \varphi d\mathbf{x} = \sum_{C \in \mathbb{T}} \int_C \mathbf{J}'_1 \cdot \nabla_x \varphi d\mathbf{x} = 0,$$

ce qui donne donc la forme non mixte

$$\boxed{\sum_{C \in \mathbb{T}} \int_C \nabla T_0^h \cdot \nabla_x \varphi d\mathbf{x} = 0 \quad \forall \varphi \in \mathbb{X}_k \cap \mathcal{C}_0^0(\Omega).} \quad (3.43)$$

### 3.5 Convergence des problèmes discrets

Les conditions aux limites fortes (3.34) et les relations (3.36), (3.37), (3.42) et (3.43) indiquent que  $\rho_0^h \in \mathbb{X}_k \cap \mathcal{C}^0(\Omega)$ ,  $(\mathbf{u}_0^h, p_1) \in (\mathbb{Y}_k \cap \mathcal{C}^0(\Omega, \mathbb{R}^3)) \times \mathbb{X}$  et  $T_0^h \in \mathbb{X}_k \cap \mathcal{C}^0(\Omega)$  sont respectivement solutions des trois problèmes :

➤ Équation de Boussinesq

$$\left\{ \begin{array}{l} \sum_{C \in \mathbb{T}} \int_C \nabla (\rho_0^h + T_0^h) \cdot \boldsymbol{\varphi} d\mathbf{x} = \int_{\Omega} \nabla (\rho_0^h + T_0^h) \cdot \boldsymbol{\varphi} d\mathbf{x} = 0 \quad \forall \boldsymbol{\varphi} \in \mathbb{Y}_k, \\ \rho_0^h|_{\Gamma} = \sigma_w(f_0^h), \end{array} \right. \quad (3.44)$$

➤ Équation de Stokes sous contrainte

$$\left\{ \begin{array}{l} \sum_{C \in \mathbb{T}} \int_C [p_1^h \nabla \cdot \boldsymbol{\varphi} - \nabla_x \mathbf{u}_0^h : \nabla_x \boldsymbol{\varphi}] d\mathbf{x} = \int_{\Omega} [p_1^h \nabla \cdot \boldsymbol{\varphi} - \nabla_x \mathbf{u}_0^h : \nabla_x \boldsymbol{\varphi}] d\mathbf{x} = 0, \\ \sum_{C \in \mathbb{T}} \int_C \nabla \cdot \mathbf{u}_0^h \varphi d\mathbf{x} = \int_{\Omega} \nabla \cdot \mathbf{u}_0^h \varphi d\mathbf{x} = 0 \\ \mathbf{u}_0^h|_{\Gamma} = \mathbf{u}_w, \end{array} \right. \quad \forall \varphi \in \mathbb{X}_k, \forall \boldsymbol{\varphi} \in \mathbb{Y}_k \cap \mathcal{C}_0^1(\Omega, \mathbb{R}^3), \quad (3.45)$$

➤ Équation de Laplace

$$\left\{ \begin{array}{l} \sum_{C \in \mathbb{T}} \int_C \nabla T_0^h \cdot \nabla_x \varphi d\mathbf{x} = \int_{\Omega} \nabla T_0^h \cdot \nabla_x \varphi d\mathbf{x} = 0 \quad \forall \varphi \in \mathbb{X}_k \cap \mathcal{C}_0^0(\Omega), \\ T_0^h|_{\Gamma} = \theta_w. \end{array} \right. \quad (3.46)$$

Il reste alors à montrer que ces problèmes sont bien posés, et que la solution de chacun de ces problèmes discrets converge bien vers celle du problème continu correspondant quand  $h \rightarrow 0$ .

### 3.5.1 Problème de Laplace

On pose :

$$V = H_0^1(\Omega), \quad V_h = \mathbb{X}_k \cap \mathcal{C}^0(\Omega), \quad V_h^0 = V_h \cap V.$$

$V_h^0$  est clairement un sous-espace vectoriel de  $V$ .

On considère le relèvement  $R_{\theta}$  vu dans la partie (3.2.1), choisi tel que  $R_{\theta} \in V_h$ . On pose  $\tilde{T}^h = T_0^h - R_{\theta} \in V_h^0$ . On a alors :

$$\begin{aligned} \sum_{C \in \mathbb{T}} \int_C \nabla \tilde{T}^h \cdot \nabla_x \varphi d\mathbf{x} &= \sum_{C \in \mathbb{T}} \int_C \nabla T_0^h \cdot \nabla_x \varphi d\mathbf{x} - \sum_{C \in \mathbb{T}} \int_C \nabla R_{\theta} \cdot \nabla_x \varphi d\mathbf{x} \\ &= \sum_{C \in \mathbb{T}} \int_C \Delta R_{\theta} \varphi d\mathbf{x} \quad \forall \varphi \in V_h^0 \end{aligned}$$

On retrouve le même problème variationnel (3.6) que dans la partie (3.2.1), défini sur un sous-espace de  $V$ . Selon un théorème classique d'estimation d'erreur pour les problèmes elliptiques (voir [16], théorème II-36, p. 57), en supposant  $\tilde{T} \in H^{m+1}(\Omega)$ ,  $m \in \{0, 1\}$ , il existe une constante  $C > 0$  telle que

$$\|T_0' - T_0^h\|_{H^1(\Omega)} = \|\tilde{T} - \tilde{T}^h\|_{H^1(\Omega)} \leq Ch^{m+1} \|\tilde{T}\|_{H^{m+1}(\Omega)}.$$

où  $h$  est la "taille de maillage" (plus abstraitement,  $h$  est un paramètre quantifiant la qualité de l'approximation de  $V$  par  $V^h$ ).

On a alors le résultat voulu :

$$\boxed{T_0^h \xrightarrow{h \rightarrow 0} T_0' \text{ dans } H^1(\Omega).} \quad (3.47)$$



### 3.5.2 Problème de Boussinesq

Comme on l'a vu au paragraphe 3.2.2, la condition  $\rho_0^h|_\Gamma = \sigma_w(f_0^h)$  est une condition de compatibilité. Compte-tenu de l'expression forte de  $f_0^h$  obtenue en (3.27), cette condition est vérifiée comme en (3.8).

On recherche la solution  $p_0^h$  de

$$c(p_0^h, \varphi) = - \int_{\Omega} \nabla p_0^h \cdot \varphi \, dx = 0 \quad \forall \varphi \in \mathbb{Y}_k. \quad (3.48)$$

On peut choisir  $\varphi = \nabla p_0^h$ , ce qui donne  $\|\nabla p_0^h\|_{L^2}^2 = 0$ ; on a donc  $p_0^h \equiv \text{cte}$ . Comme dans le cas continu, cette constante est indéterminée a priori; le choix de chercher une solution dans l'espace  $L_0^2$  impose alors

$$p_0^h \equiv 0, \quad (3.49)$$

et l'on a donc  $\rho_0^h + T_0^h = p_0^h = p_0 = \rho_0 + T_0$  pour tout  $h$ , soit avec (3.47) :

$$\boxed{\rho_0^h \xrightarrow{h \rightarrow 0} \rho'_0 \text{ dans } H^1(\Omega).} \quad (3.50)$$

### 3.5.3 Problème de Stokes

La discrétisation Galerkin Discontinu du problème de Stokes stationnaire sous contraintes est étudiée de façon précise par Di Pietro et Ern dans [59]. Il y est montré que la résolution d'un problème mixte de type point-selle comme celui de Stokes, qui se met sous la forme

$$\begin{cases} a(\mathbf{u}_0^h, \varphi) + b(\varphi, p_1^h) = L(\varphi) & \forall \varphi \in \mathbb{Y}^h \\ -b(\mathbf{u}^h, \psi) = G(\psi) & \forall \psi \in \mathbb{X}^h, \end{cases} \quad (3.51)$$

avec  $\mathbb{X}^h \subset L_0^2(\Omega)$  et  $\mathbb{Y}^h \subset H_0^1(\Omega, \mathbb{R}^3)$  deux espaces d'approximation de "pas de maillage"  $h$ , respectivement pour la pression et la vitesse, et les formes bilinéaires  $a$  et  $b$  définies comme dans la partie 3.2.3 par

$$\begin{cases} a : (\mathbf{u}, \varphi) \in H_0^1(\Omega, \mathbb{R}^3) \times H_0^1(\Omega, \mathbb{R}^3) \mapsto \int_{\Omega} \nabla \mathbf{u} : \nabla \varphi \, dx, \\ b : (\mathbf{u}, \psi) \in H_0^1(\Omega, \mathbb{R}^3) \times L_0^2(\Omega) \mapsto - \int_{\Omega} \psi \nabla \cdot \mathbf{u} \, dx \end{cases}$$

n'est bien posé que si  $b$  vérifie la condition inf-sup sur  $\mathbb{Y}^h \times \mathbb{X}^h$

$$\exists \beta_h > 0 / \inf_{\psi \in \mathbb{X}^h \setminus \{0\}} \sup_{\varphi \in \mathbb{Y}^h \setminus \{0\}} \frac{b(\varphi, \psi)}{\|\varphi\|_{H_0^1} \|\psi\|_{L^2}} \geq \beta_h, \quad (3.52)$$

et que dans ce cas,  $(\mathbf{u}_0^h, p_1^h) \xrightarrow{h \rightarrow 0} (\mathbf{u}'_0, p'_1)$  si et seulement si

➤ la condition inf-sup vérifiée par  $b$  est uniforme :

$$\exists \beta > 0 / \forall h > 0, \beta_h \geq \beta$$

➤ les espaces d'approximation sont assez fins :

$$\begin{cases} \forall \mathbf{u} \in H_0^1(\Omega, \mathbb{R}^3), d(\mathbf{u}, \mathbb{Y}^h) = \inf_{\boldsymbol{\varphi} \in \mathbb{Y}^h} \|\mathbf{u} - \boldsymbol{\varphi}\| \xrightarrow{h \rightarrow 0} 0, \\ \forall p \in L_0^2(\Omega), d(p, \mathbb{X}^h) = \inf_{\psi \in \mathbb{X}^h} \|p - \psi\| \xrightarrow{h \rightarrow 0} 0. \end{cases}$$

☞ NB : la coercivité de la forme bilinéaire  $a$  est quant à elle naturellement préservée par la restriction de  $H_0^1(\Omega, \mathbb{R}^3)$  à  $\mathbb{Y}^h$  (qui ne peut même que la rendre plus forte). La continuité des formes bilinéaires est bien entendu également conservée.  $\square$

Une discrétisation simple comme celle que nous obtenons et qui a permis d'aboutir à un résultat de convergence dans le cas du problème de Laplace précédent, consistant à chercher une approximation de la solution  $(\mathbf{u}_0', p_1') \in H_0^1(\Omega, \mathbb{R}^3) \times L_0^2(\Omega)$  du problème continu dans les espaces d'approximation

$$\begin{cases} \mathbb{X}^h = \mathbb{X}_k \cap L_0^2(\Omega) \\ \mathbb{Y}^h = \mathbb{Y}_k \cap H_0^1(\Omega, \mathbb{R}^3) \end{cases}$$

est vouée à l'échec. Ceci est dû à la mise en défaut de la condition inf-sup (3.52) lorsqu'on choisit dans les espaces d'approximation des polynômes de degré identique pour la vitesse et la pression, ce qui traduit une perte de contrôle des sauts de pression aux interfaces du maillage. En effet, cette condition traduit simplement la surjectivité de l'opérateur de divergence

$$\begin{aligned} D : \mathbb{Y} &\rightarrow \mathbb{X} \\ \mathbf{v} &\mapsto \nabla \cdot \mathbf{v} \end{aligned}$$

qui est vérifiée pour  $\mathbb{Y} = H_0^1(\Omega, \mathbb{R}^3)$  et  $\mathbb{X} = L_0^2(\Omega)$  (lemme de Nečas), mais qui ne l'est plus dans leurs restrictions  $\mathbb{Y}^h$  et  $\mathbb{X}^h$  dès lors que  $\dim(\mathbb{Y}^h) \leq \dim(\mathbb{X}^h)$ .

Dans le cas d'une approximation par polynômes d'ordre 1 pour la vitesse et la pression vient s'ajouter le phénomène de "verrouillage numérique" de la vitesse (locking) : le noyau à gauche de la forme  $b$  est réduit à  $\{\mathbf{0}\}$  :

$$\{\boldsymbol{\varphi} \in \mathbb{Y}^h, b(\boldsymbol{\varphi}, \psi) = 0 \quad \forall \psi \in \mathbb{X}\} = \{\mathbf{0}\}. \quad (3.53)$$

Ce phénomène est dû à une contrainte d'incompressibilité trop forte.

Il existe deux possibilités pour obtenir une discrétisation correcte de ce problème :

- utiliser des polynômes de degrés différents pour la vitesse et la pression, typiquement des polynômes de degré 2 pour la vitesse et de degré 1 pour la pression ;
- utiliser des polynômes de même degré en introduisant dans la formulation variationnelle discrète des termes de stabilisation visant à contrôler les sauts de pression tout en préservant la consistance du problème discret (c'est la méthode dite de "compressibilité artificielle", où l'on réintroduit l'inconnue de pression dans la contrainte d'incompressibilité).

La première solution ne nous est pas accessible, la discrétisation de  $\mathbf{u}_0'$  et  $p_1'$  étant directement héritée de celle de la distribution  $f$  : il n'y a qu'une seule variable au niveau cinétique, les espaces d'approximation  $\mathbb{X}^h$  et  $\mathbb{Y}^h$  sont nécessairement construits sur la même base polynomiale.

La seconde solution, introduite dans [26], consiste à mettre le problème variationnel (3.51) sous la forme

$$\begin{cases} \tilde{a}(\mathbf{u}_0^h, \varphi) + \tilde{b}(\varphi, p_1^h) = \tilde{L}(\varphi) & \forall \varphi \in \mathbb{Y} \\ -\tilde{b}(\mathbf{u}_0^h, \psi) + \tilde{c}(\psi, p_1^h) = \tilde{G}(\psi) & \forall \psi \in \mathbb{X}, \end{cases}$$

où  $\tilde{c}(\psi, p_1^h)$  est un terme de stabilisation des sauts de pression. Di Pietro et Ern proposent par exemple dans [59] la formulation variationnelle

$$\begin{cases} \tilde{a}(\mathbf{u}_0^h, \varphi) = \int_{\Omega} \nabla_x \mathbf{u}_0^h : \nabla_x \varphi \, d\mathbf{x} + \sum_{K \in \mathbb{F}} \int_K \left[ \frac{\eta}{h_K} \llbracket \mathbf{u}_{0i}^h \rrbracket \llbracket \varphi_i \rrbracket - (\llbracket \mathbf{u}_{0i}^h \rrbracket \{ \nabla \varphi_i \} + \llbracket \varphi_i \rrbracket \{ \nabla \mathbf{u}_{0i}^h \}) \cdot \mathbf{n} \right] d\mathbf{x}, \\ \tilde{b}(\mathbf{u}_0^h, \psi) = - \int_{\Omega} \psi \nabla \cdot \mathbf{u}_0^h \, d\mathbf{x} + \sum_{K \in \mathbb{F}} \int_K \{ \psi \} \llbracket \mathbf{u}_0^h \rrbracket \cdot \mathbf{n} \, d\mathbf{x}, \\ \tilde{c}(\psi, p_1^h) = \sum_{K \in \mathring{\mathbb{F}}} \int_K h_K \llbracket \psi \rrbracket \llbracket p_1^h \rrbracket \, d\mathbf{x}, \end{cases}$$

où  $\eta > 0$  est un paramètre arbitraire et  $h_K$  une échelle de longueur locale associée à l'arête  $K$ .

Il ne nous est pas possible de faire apparaître ces termes dans le cadre de notre étude : l'équation d'incompressibilité en particulier est obtenue sans faire la moindre restriction sur les fonctions-test, ce qui ne laisse aucune possibilité de faire apparaître le terme  $\tilde{c}$ .

Dans le cas particulier évoqué précédemment où  $\mathbf{u}_w$  est constant, le problème (3.45) admet pour solution triviale  $(\mathbf{u}_0^h, p_1^h) \equiv (\mathbf{u}_w, 0)$ , ce qui correspond à la solution triviale du problème continu avec la même hypothèse. Le verrouillage numérique montre alors que dans ce cas particulier, le problème est bien posé avec une discrétisation  $\mathbb{P}_1/\mathbb{P}_1$  : en effectuant le relèvement  $\mathbf{R}_u \equiv \mathbf{u}_w$ , soit  $\mathbf{u}_0^h = \tilde{\mathbf{u}}^h + \mathbf{u}_w$ , on obtient le problème discret

$$\begin{cases} a(\tilde{\mathbf{u}}^h, \varphi) + b(\varphi, p_1^h) = 0 & \forall \varphi \in \mathbb{Y}^h, \\ -b(\tilde{\mathbf{u}}^h, \psi) = - \int_{\Omega} \nabla \cdot \mathbf{u}_w \psi \, d\mathbf{x} = 0 & \forall \psi \in \mathbb{X}^h; \end{cases}$$

la seconde équation donne avec (3.53) :  $\tilde{\mathbf{u}}^h \equiv \mathbf{0}$ , donc  $\mathbf{u}_0^h \equiv \mathbf{u}_w$ .

### 3.6 Analyse asymptotique : termes d'ordre $n \geq 2$

On rappelle l'équation (3.25) :

$$\begin{aligned} \forall g \in \mathbb{E}_k, \quad & \sum_{C \in \mathbb{T}} \int_C \left\langle f_n^h \mathbf{v} \cdot \nabla_x g M^{-1} + L(f_{n+1}^h) g M^{-1} \right\rangle d\mathbf{x} - \sum_{K \in \mathring{\mathbb{F}}} \int_K \left\langle \mathbf{v} \cdot \tilde{\mathbf{n}} \llbracket f_n^h \rrbracket \llbracket g \rrbracket M^{-1} \right\rangle d\mathbf{x} \\ & = - \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle f_n^h g M^{-1} (\mathbf{v} \cdot \mathbf{n})^- \right\rangle d\mathbf{x} + \sqrt{2\pi} \sum_{K \in \mathbb{F}^\partial} \int_K \left\langle (\mathbf{v} \cdot \mathbf{n})^- f_n^h \right\rangle \left\langle (\mathbf{v} \cdot \mathbf{n})^+ g \right\rangle d\mathbf{x}. \end{aligned}$$

De la même façon que pour l'ordre 1, le choix de fonction-test  $g = \varphi, \varphi M$ ,  $\varphi \in \mathbb{X}_k \cap \mathcal{C}^0(\Omega)$  donne la continuité à l'ordre  $n$  :

$$\boxed{\sum_{C \in \mathbb{T}} \int_C \mathbf{u}_n^h \cdot \nabla_x \varphi d\mathbf{x} = 0 \quad \forall \varphi \in \mathbb{X}_k \cap \mathcal{C}_0^0(\Omega)} \quad (3.54)$$

En revanche, les équations de Stokes ou de Laplace deviennent bien plus complexes à obtenir. Avec la fonction-test  $g = \mathbf{v} \cdot \boldsymbol{\varphi} M$ ,  $\boldsymbol{\varphi} \in \mathbb{Y}_k \cap \mathcal{C}_0^0(\Omega)$ , on obtient

$$\sum_{C \in \mathbb{T}} \int_C \bar{\bar{J}}_n : \nabla_x \boldsymbol{\varphi} d\mathbf{x} = 0 \quad \forall \boldsymbol{\varphi} \in \mathbb{Y} \cap \mathcal{C}^0(\Omega), \quad \bar{\bar{J}}_n = \left\langle (\mathbf{v} \otimes \mathbf{v}) f_n^h \right\rangle, \quad (3.55)$$

mais on a ensuite besoin pour établir une formulation mixte de l'équivalent de la relation (3.38) à l'ordre  $n$ , c'est-à-dire de la relation (3.25) à l'ordre  $n-1$ . Or à moins de restreindre grossièrement l'espace des fonctions-tests, au risque d'obtenir un problème mal posé, les calculs deviennent lourds : la fonction de distribution  $f_{n-1}^h$  n'est *a priori* pas continue, et ses valeurs au bord sont inconnues. Le calcul des termes sur  $\mathbb{F}^\partial$  est donc complexe. On arrêtera notre étude ici.

### 3.7 Bilan

Après avoir établi les trois problèmes aux limites permettant d'obtenir les grandeurs  $\rho'_0, \mathbf{u}_0'$  et  $T'_0$ , respectivement de type Boussinesq, Stokes et Laplace, nous avons démontré que ces problèmes sont bien posés sous la contrainte

$$\int_{\Omega} p'_0 d\mathbf{x} = \int_{\Omega} (\rho'_0 + T'_0) d\mathbf{x} = 0,$$

qui fixe l'amplitude de la perturbation de l'écoulement.

Nous avons ensuite construit un schéma Galerkin Discontinu avec un flux numérique "upwind", que nous avons appliqué au problème BGK linéarisé avec condition de réflexion diffuse étudié au chapitre précédent, obtenant ainsi une approximation  $f^h$  de la fonction de distribution  $f$ . L'application du développement de Hilbert à cette distribution discrète  $f^h$  a permis d'obtenir à tout ordre  $n \in \mathbb{N}$  du développement un problème discret déterminant une approximation  $f_n^h$  de  $f_n$ . Le choix à chaque ordre de fonctions-test adaptées dans le schéma a permis d'établir

- à l'ordre  $-1$  une expression forte de  $f_0^h$  sous forme de maxwellienne linéarisée ;
- à l'ordre 0
  - la continuité de la fonction de distribution discrète  $f_0^h$ , en faisant l'hypothèse que  $\theta_w$  est uniforme,
  - sous la même hypothèse, les conditions aux limites d'adhérence sur les grandeurs macroscopiques qui en sont issues :

$$\begin{cases} \mathbf{u}_0^h = \mathbf{u}_w, \\ T_0^h = \theta_w. \end{cases}$$

ainsi que la condition de compatibilité  $\rho_0^h = \sigma_w(f_0^h)$  ;

- une équation de continuité sur  $\mathbf{u}_0^h$  et une équation de Boussinesq sur  $p_0^h = \rho_0^h + T_0^h$  ;
- à l'ordre 1
  - une équation de continuité sur  $\mathbf{u}_1^h$
  - une équation de Stokes sur  $\mathbf{u}_0^h$  et  $p_1^h = \rho_1^h + T_1^h$  et une équation de Laplace sur  $T_0^h$  sous forme mixte, que nous avons ensuite exprimées sous forme simple moyennant des hypothèses de régularité supplémentaires.

Nous avons enfin analysé les problèmes de Laplace, Stokes et Boussinesq discrets obtenus. Les problèmes de Laplace et de Boussinesq sont bien posés, et leur solution respective converge inconditionnellement vers la solution du problème continu correspondant dans  $H^1(\Omega)$ . En revanche, le problèmes de Stokes est plus délicat à traiter. Le fait d'être contraint à utiliser la même discrétisation pour toutes les variables (celles-ci étant toutes des moments de  $f$  par rapport à la variable  $\mathbf{v}$ ) fait que ce problème discret est mal posés dans le cas général. Dans le cas particulier où les perturbations à la paroi  $\mathbf{u}_w$  et  $\theta_w$  sont uniformes, le problème continu admet une solution triviale que le schéma Galerkin Discontinu est capable de capturer.

Enfin, l'analyse de l'ordre 1, où on espère obtenir des problèmes de Stokes et Laplace avec des conditions de glissement, est très complexe et n'a pas été faite ici.

Dans le cas général, les schémas Galerkin Discontinu ne présentent donc pas de propriété de préservation asymptotique complète pour le problème BGK linéarisé avec réflexion diffuse. Cependant, ces schémas restent performants et peuvent montrer des atouts non négligeables face aux schémas classiques comme les schémas Volumes Finis. Le chapitre suivant présente une étude numérique comparative entre ces deux types de schémas appliqués à l'équation BGK pour un problème de Couette qui le met en évidence.

## Chapitre 4

# Résolution numérique de l'équation BGK 1D : écoulement de Couette

*L'unique moyen de savoir jusqu'où l'on peut aller, c'est de se mettre en route et de marcher.*

Henri Bergson - *L'Energie spirituelle*

On présente dans ce chapitre une étude numérique comparant les performances de quelques schémas Volumes Finis d'ordres divers (allant de l'ordre 1 à l'ordre 4) avec un schéma Galerkin Discontinu simple sur l'équation BGK avec conditions aux limites de réflexion diffuse. Dans un souci de simplicité, on se place dans un cas monodimensionnel en espace : le cas-test considéré est un écoulement de type Couette de gaz monoatomique entre deux plaques planes.

Dans un premier temps, on présente un modèle réduit à trois distributions marginales permettant de calculer cet écoulement avec un espace de vitesses 1D. On décrit ensuite le schéma à vitesses discrètes générique, c'est-à-dire semi-discret en espace, appliqué à ce modèle, permettant une implication du terme de collision grâce à une linéarisation de l'équilibre. Enfin, on détaille l'utilisation de ce schéma avec des méthodes Volumes Finis et avec une méthode Galerkin Discontinu, en exposant la discrétisation des termes de transport et en définissant des conditions aux limites appropriées à cette discrétisation. On compare enfin les résultats obtenus sur différents maillages.

### 4.1 Modèle à distributions marginales

On cherche à modéliser un écoulement de Couette monodimensionnel de gaz monoatomique. La configuration de l'écoulement est donnée dans la figure 4.1 : le gaz s'écoule entre deux plaques planes, de même température, animées chacune d'une vitesse purement tangentielle ( $\mathbf{u}_w = u_w \mathbf{y}$ ).

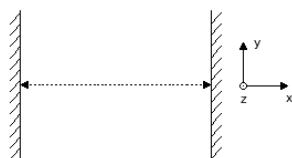


FIGURE 4.1: Configuration de l'écoulement de Couette

L'écoulement étant 1D, on a  $\partial_y = \partial_z = 0$ . En notant  $\mathbf{v} = (v_x, v_y, v_z)^T$  la vitesse mésoscopique, l'équation BGK se réduit alors à :

$$\frac{\partial F}{\partial t} + v_x \frac{\partial F}{\partial x} = \frac{1}{\tau} (\mathcal{M}[\mathbf{U}] - F), \quad F(t, \mathbf{x}, \mathbf{v}) \quad (4.1)$$

où  $\mathbf{U}$  est le vecteur des quantités conservatives (densité, quantités de mouvement et énergie totale). En notant  $\rho$  la densité,  $\mathbf{u} = (u, v, w)^T$ , la vitesse macroscopique et  $e$  la densité d'énergie totale, ce vecteur s'écrit

$$\mathbf{U}(t, \mathbf{x}) = \begin{pmatrix} \rho(t, \mathbf{x}) \\ \rho u(t, \mathbf{x}) \\ \rho v(t, \mathbf{x}) \\ \rho e(t, \mathbf{x}) \end{pmatrix} = \int_{\mathbb{R}^3} \mathbf{m}(\mathbf{v}) F(t, \mathbf{x}, v_x, v_y, v_z) dv_x dv_y dv_z, \quad \mathbf{m}(\mathbf{v}) = \left( 1, v_x, v_y, \frac{1}{2} |\mathbf{v}|^2 \right)^T$$

avec  $|\mathbf{v}|^2 = v_x^2 + v_y^2 + v_z^2$ . On ne considère pas la quantité de mouvement dans la direction  $z$ , la géométrie de l'écoulement donnant immédiatement  $w = 0$ .

La maxwellienne d'équilibre  $\mathcal{M}[\mathbf{U}]$  s'écrit en fonction de la densité, de la vitesse macroscopique et de la température  $T$  :

$$\mathcal{M}[\mathbf{U}] = \mathcal{M}[\rho, \mathbf{u}, T] = \frac{\rho}{(2\pi RT)^{3/2}} \exp \left( -\frac{|u - v_x|^2 + |v - v_y|^2 + v_z^2}{2RT} \right).$$

Afin de simplifier notre problème, on utilise un modèle à distributions marginales pour résoudre l'équation (4.1). Cette technique, introduite par Chu en 1965 [25], consiste à éliminer une ou deux variables de vitesse en intégrant l'équation par rapport à ces variables, puis en introduisant des fonctions de distribution marginales correspondant aux moments nécessaires pour exprimer les grandeurs macroscopiques. On écrit alors une équation d'évolution pour chacune de ces distributions marginales.

On souhaite ici ne conserver que la vitesse  $v_x$  : on considère pour cela les distributions marginales

$$\begin{cases} f(t, \mathbf{x}, v_x) = \int_{\mathbb{R}^2} F(t, \mathbf{x}, v_x, v_y, v_z) dv_y dv_z, \\ g(t, \mathbf{x}, v_x) = \int_{\mathbb{R}^2} \frac{1}{2} (v_y^2 + v_z^2) F(t, \mathbf{x}, v_x, v_y, v_z) dv_y dv_z, \\ h(t, \mathbf{x}, v_x) = \int_{\mathbb{R}^2} v_y F(t, \mathbf{x}, v_x, v_y, v_z) dv_y dv_z. \end{cases} \quad (4.2)$$

Les quantités macroscopiques conservatives s'écrivent alors :

$$\begin{cases} \rho = \int_{\mathbb{R}^3} F dv_x dv_y dv_z = \int_{\mathbb{R}} f dv_x, \\ \rho u = \int_{\mathbb{R}^3} v_x F dv_x dv_y dv_z = \int_{\mathbb{R}} v_x f dv_x, \\ \rho v = \int_{\mathbb{R}^3} v_y F dv_x dv_y dv_z = \int_{\mathbb{R}} h dv_x, \\ \rho e = \int_{\mathbb{R}^3} \frac{1}{2} (v_x^2 + v_y^2 + v_z^2) F dv_x dv_y dv_z = \int_{\mathbb{R}} \frac{1}{2} v_x^2 f dv_x + \int_{\mathbb{R}} g dv_x. \end{cases}$$

Le flux de chaleur s'exprime :

$$\begin{aligned}
q_x &= \int_{\mathbb{R}^3} \frac{1}{2} |\mathbf{u} - \mathbf{v}|^2 (u - v_x) F dv_x dv_y dv_z \\
&= \int_{\mathbb{R}^3} \frac{1}{2} (|u - v_x|^2 + |v - v_y|^2 + v_z^2) (u - v_x) F dv_x dv_y dv_z \\
&= \int_{\mathbb{R}^3} \frac{1}{2} |u - v_x|^2 (u - v_x) F dv_x dv_y dv_z + \int_{\mathbb{R}^3} \frac{1}{2} (v_y^2 + v_z^2) (u - v_x) F dv_x dv_y dv_z \\
&\quad - 2v \int_{\mathbb{R}^3} \frac{1}{2} v_y (u - v_x) F dv_x dv_y dv_z + \underbrace{v^2 \int_{\mathbb{R}^3} \frac{1}{2} (u - v_x) F dv_x dv_y dv_z}_{=0} \\
&= \int_{\mathbb{R}} \frac{1}{2} (u - v_x)^3 f dv_x + \int_{\mathbb{R}} (u - v_x) g dv_x - v \int_{\mathbb{R}} (u - v_x) h dv_x \\
&= \int_{\mathbb{R}} \frac{1}{2} (u - v_x)^3 f dv_x + \int_{\mathbb{R}} (u - v_x) (g - v h) dv_x.
\end{aligned}$$

et l'énergie totale s'écrit :

$$\rho e = \frac{1}{2} \rho \mathbf{u}^2 + \frac{3}{2} \rho R T = \frac{1}{2} \rho (u^2 + v^2) + \frac{3}{2} \rho R T,$$

ce qui donne la température et la pression :

$$\begin{cases} T = \frac{2}{3R} \left( e - \frac{1}{2} (u^2 + v^2) \right), \\ p = \rho R T = \frac{2}{3} \rho \left( e - \frac{1}{2} (u^2 + v^2) \right). \end{cases} \quad (4.3)$$

On voit donc apparaître un modèle à 3 distributions,  $f$ ,  $g$  et  $h$ , ne dépendant que de la position et de la vitesse  $v_x$ . L'espace des vitesses discrètes correspondant sera noté  $\mathcal{V}$  :

$$\mathcal{V} = \{v_k\}_{k \in \mathcal{K} \subset \mathbb{N}}.$$

On doit maintenant écrire une équation d'évolution pour chacune des trois distributions réduites du modèle.

#### 4.1.1 Équation pour la distribution $f$

La distribution réduite  $f$  s'obtenant par intégration de la distribution initiale  $F$  selon (4.2), on obtient une équation sur  $f$  en intégrant de la même façon l'équation initiale (4.1), ce qui donne

$$\frac{\partial f}{\partial t} + v_x \frac{\partial f}{\partial x} = \frac{1}{\tau} \left( \int_{\mathbb{R}^2} \mathcal{M}[\mathbf{U}] dv_y dv_z - f \right).$$

**Proposition 5.**  $\mathcal{M}[\mathbf{U}] := \int_{\mathbb{R}^2} \mathcal{M}[\mathbf{U}] dv_y dv_z = \frac{\rho}{\sqrt{2\pi RT}} \exp\left(-\frac{|u - v_x|^2}{2RT}\right).$



*Démonstration 5.*  $\int_{\mathbb{R}} \mathcal{M}[\mathbf{U}] dv_z = \frac{\rho}{(2\pi RT)^{\frac{3}{2}}} \exp\left(-\frac{|u-v_x|^2 + |v-v_y|^2}{2RT}\right) \int_{\mathbb{R}} \exp\left(-\frac{v_z^2}{2RT}\right) dv_z,$   
or  $\int_{\mathbb{R}} \exp\left(-\frac{v_z^2}{2RT}\right) dv_z = \left[\frac{\sqrt{2\pi RT}}{2} \operatorname{erf}\left(\frac{v_z}{\sqrt{2RT}}\right)\right]_{-\infty}^{+\infty} = \sqrt{2\pi RT},$   
donc  $\int_{\mathbb{R}} \mathcal{M}[\mathbf{U}] dv_z = \frac{\rho}{2\pi RT} \exp\left(-\frac{|u-v_x|^2 + |v-v_y|^2}{2RT}\right) = \mathfrak{M}[\mathbf{U}],$   
et  $\int_{\mathbb{R}^2} \mathcal{M}[\mathbf{U}] dv_y dv_z = \int_{\mathbb{R}} \frac{\rho}{2\pi RT} \exp\left(-\frac{|u-v_x|^2 + |v-v_y|^2}{2RT}\right) dv_y$   

$$= \frac{\rho}{2\pi RT} \exp\left(-\frac{|u-v_x|^2}{2RT}\right) \int_{\mathbb{R}} \exp\left(-\frac{|v-v_y|^2}{2RT}\right) dv_y,$$
  
or  $\int_{\mathbb{R}} \exp\left(-\frac{|v-v_y|^2}{2RT}\right) dv_y = \int_{\mathbb{R}} \exp\left(-\frac{z^2}{2RT}\right) dz = \sqrt{2\pi RT} \quad (z = v - v_y),$   
donc  $\int_{\mathbb{R}^2} \mathcal{M}[\mathbf{U}] dv_y dv_z = \int_{\mathbb{R}} \mathfrak{M}[\mathbf{U}] dv_y = \frac{\rho}{\sqrt{2\pi RT}} \exp\left(-\frac{|u-v_x|^2}{2RT}\right). \quad \square$

#### 4.1.2 Équation pour la distribution $g$

On procède pour la fonction  $g$  de la même façon que pour  $f$ , en intégrant l'équation (4.1) avec la mesure adaptée, ce qui donne la relation

$$\frac{\partial g}{\partial t} + v_x \frac{\partial g}{\partial x} = \frac{1}{\tau} \left( \int_{\mathbb{R}^2} \frac{1}{2} (v_y^2 + v_z^2) \mathcal{M}[\mathbf{U}] dv_y dv_z - g \right).$$

**Proposition 6.**  $\mathcal{N}[\mathbf{U}] := \int_{\mathbb{R}^2} \frac{1}{2} (v_y^2 + v_z^2) \mathcal{M}[\mathbf{U}] dv_y dv_z = \left( RT + \frac{v^2}{2} \right) \mathcal{M}[\mathbf{U}].$

*Démonstration 6.* On commence par calculer  $\int_{\mathbb{R}} \frac{1}{2} v_z^2 \mathcal{M}[\mathbf{U}] dv_z :$

$$\int_{\mathbb{R}} \frac{1}{2} v_z^2 \mathcal{M}[\mathbf{U}] dv_z = \frac{1}{2} \frac{\rho}{(2\pi RT)^{\frac{3}{2}}} \exp\left(-\frac{|u-v_x|^2 + |v-v_y|^2}{2RT}\right) \int_{\mathbb{R}} v_z^2 \exp\left(-\frac{v_z^2}{2RT}\right) dv_z.$$

En intégrant par parties :

$$u' = v_z \exp\left(-\frac{v_z^2}{2RT}\right) \Rightarrow u = -RT \exp\left(-\frac{v_z^2}{2RT}\right), \quad v = v_z \Rightarrow v' = 1., \text{ et on obtient}$$

$$\int_{\mathbb{R}} v_z^2 \exp\left(-\frac{v_z^2}{2RT}\right) dv_z = \underbrace{\left[-v_z RT \exp\left(-\frac{v_z^2}{2RT}\right)\right]_{-\infty}^{+\infty}}_{=0} + \underbrace{RT \int_{\mathbb{R}} \exp\left(-\frac{v_z^2}{2RT}\right) dv_z}_{=\sqrt{2\pi RT}} = RT \sqrt{2\pi RT},$$

$$\text{d'où : } \int_{\mathbb{R}} \frac{1}{2} v_z^2 \mathcal{M}[\mathbf{U}] dv_z = \frac{1}{2} \frac{\rho}{(2\pi RT)^{\frac{3}{2}}} \exp\left(-\frac{|u-v_x|^2 + |v-v_y|^2}{2RT}\right) RT \sqrt{2\pi RT}$$

$$\begin{aligned}
&= \frac{RT}{2} \frac{\rho}{2\pi RT} \exp\left(-\frac{|u - v_x|^2 + |v - v_y|^2}{2RT}\right) \\
&= \frac{RT}{2} \mathfrak{M}[\mathbf{U}],
\end{aligned}$$

et donc :  $\int_{\mathbb{R}^2} \frac{1}{2} v_z^2 \mathcal{M}[\mathbf{U}] dv_y dv_z = \int_{\mathbb{R}} \frac{RT}{2} \mathfrak{M}[\mathbf{U}] dv_y = \frac{RT}{2} \mathcal{M}[\mathbf{U}].$

De même :  $\int_{\mathbb{R}} \frac{1}{2} v_y^2 \mathcal{M}[\mathbf{U}] dv_y = \frac{1}{2} \frac{\rho}{(2\pi RT)^{\frac{3}{2}}} \exp\left(-\frac{|u - v_x|^2 + v_z^2}{2RT}\right) \int_{\mathbb{R}} v_y^2 \exp\left(-\frac{|v - v_y|^2}{2RT}\right) dv_y.$

Or  $\int_{\mathbb{R}} v_y^2 \exp\left(-\frac{|v - v_y|^2}{2RT}\right) dv_y = \int_{\mathbb{R}} (z + v)^2 \exp\left(-\frac{z^2}{2RT}\right) dz \quad (z = v_y - v)$

$$\begin{aligned}
&= \underbrace{\int_{\mathbb{R}} z^2 \exp\left(-\frac{z^2}{2RT}\right) dz}_{= RT\sqrt{2\pi RT}} + 2v \underbrace{\int_{\mathbb{R}} z \exp\left(-\frac{z^2}{2RT}\right) dz}_{= 0} + v^2 \underbrace{\int_{\mathbb{R}} \exp\left(-\frac{z^2}{2RT}\right) dz}_{= \sqrt{2\pi RT}} \\
&= \sqrt{2\pi RT}(RT + v^2),
\end{aligned}$$

donc  $\int_{\mathbb{R}} \frac{1}{2} v_y^2 \mathcal{M}[\mathbf{U}] dv_y = \frac{1}{2} (RT + v^2) \frac{\rho}{2\pi RT} \exp\left(-\frac{|u - v_x|^2 + v_z^2}{2RT}\right),$

et  $\int_{\mathbb{R}^2} \frac{1}{2} v_y^2 \mathcal{M}[\mathbf{U}] dv_y dv_z = \frac{1}{2} (RT + v^2) \frac{\rho}{2\pi RT} \exp\left(-\frac{|u - v_x|^2}{2RT}\right) \int_{\mathbb{R}} \exp\left(-\frac{v_z^2}{2RT}\right) dv_z$

$$= \frac{1}{2} (RT + v^2) \frac{\rho}{\sqrt{2\pi RT}} \exp\left(-\frac{|u - v_x|^2}{2RT}\right) = \frac{1}{2} (RT + v^2) \mathcal{M}[\mathbf{U}].$$

On obtient donc finalement :

$$\int_{\mathbb{R}^2} \frac{1}{2} (v_y^2 + v_z^2) \mathcal{M}[\mathbf{U}] dv_y dv_z = \frac{1}{2} (RT + v^2) \mathcal{M}[\mathbf{U}] + \frac{RT}{2} \mathcal{M}[\mathbf{U}] = \left(RT + \frac{v^2}{2}\right) \mathcal{M}[\mathbf{U}]. \quad \square$$

### 4.1.3 Équation pour la distribution $h$

Enfin, toujours par le même processus, on obtient l'équation régissant l'évolution de  $h$  :

$$\frac{\partial h}{\partial t} + v_x \frac{\partial h}{\partial x} = \frac{1}{\tau} \left( \int_{\mathbb{R}^2} v_y \mathcal{M}[\mathbf{U}] dv_y dv_z - h \right).$$

**Proposition 7.**  $\mathcal{P}[\mathbf{U}] := \int_{\mathbb{R}^2} v_y \mathcal{M}[\mathbf{U}] dv_y dv_z = v \mathcal{M}[\mathbf{U}].$

*Démonstration 7.*  $\int_{\mathbb{R}} v_y \mathcal{M}[\mathbf{U}] dv_y = \frac{\rho}{(2\pi RT)^{\frac{3}{2}}} \exp\left(-\frac{|u - v_x|^2 + v_z^2}{2RT}\right) \int_{\mathbb{R}} v_y \exp\left(-\frac{|v_y - v|^2}{2RT}\right) dv_y,$

et

$$\begin{aligned}
 \int_{\mathbb{R}} v_y \exp\left(-\frac{|v_y - v|^2}{2RT}\right) dv_y &= \int_{\mathbb{R}} (z + v) \exp\left(-\frac{z^2}{2RT}\right) dz \\
 &= \underbrace{\int_{\mathbb{R}} z \exp\left(-\frac{z^2}{2RT}\right) dz}_{=0} + v \underbrace{\int_{\mathbb{R}} \exp\left(-\frac{z^2}{2RT}\right) dz}_{=\sqrt{2\pi RT}} \\
 &= v \sqrt{2\pi RT},
 \end{aligned}$$

$$\text{donc } \int_{\mathbb{R}} v_y \mathcal{M}[\mathbf{U}] dv_y = v \frac{\rho}{2\pi RT} \exp\left(-\frac{|u - v_x|^2 + v_z^2}{2RT}\right),$$

$$\begin{aligned}
 \text{et } \int_{\mathbb{R}^2} v_y \mathcal{M}[\mathbf{U}] dv_y dv_z &= v \frac{\rho}{2\pi RT} \exp\left(-\frac{|u - v_x|^2}{2RT}\right) \int_{\mathbb{R}} \exp\left(-\frac{v_z^2}{2RT}\right) dv_z \\
 &= v \frac{\rho}{\sqrt{2\pi RT}} \exp\left(-\frac{|u - v_x|^2}{2RT}\right) = v \mathcal{M}[\mathbf{U}]. \quad \square
 \end{aligned}$$

#### 4.1.4 Conditions aux limites

La condition de réflexion diffuse imposée à  $F$  se compose de deux conditions distinctes :

- un flux de masse nul à travers la paroi,
- une adaptation à la vitesse et à la température de cette même paroi ( $u_w$  et  $T_w$ ) via un équilibre,

ce qui donne

$$F = \sigma(F) \mathcal{M}[\mathbf{U}_{\mathbf{w}}],$$

avec  $\mathbf{U}_{\mathbf{w}}$  le vecteur des quantités conservatives évalué à la paroi

$$\mathbf{U}_{\mathbf{w}} = \left(1, 0, u_w, \frac{3}{2}RT_w + \frac{1}{2}u_w^2\right)^T,$$

et  $\sigma(F)$  le coefficient de réflexion défini pour assurer la condition de flux nul, qui s'écrit compte-tenu de la géométrie du domaine de calcul

$$\begin{aligned}
 \sigma(F) &= - \left( \int_{\mathbb{R}^3} v_x^- F dv_x dv_y dv_z \right) \left( \int_{\mathbb{R}^3} v_x^+ \mathcal{M}[\mathbf{U}_{\mathbf{w}}] dv_x dv_y dv_z \right)^{-1} \\
 &= - \left( \int_{\mathbb{R}} v_x^- f dv_x \right) \left( \int_{\mathbb{R}} v_x^+ \mathcal{M}[\mathbf{U}_{\mathbf{w}}] dv_x \right)^{-1} = \sigma(f);
 \end{aligned}$$

la première condition se répercute donc entièrement sur la distribution  $f$ .

On obtient alors très simplement

$$F = \sigma(F) \mathcal{M}[\mathbf{U}_{\mathbf{w}}] = \sigma(f) \mathcal{M}[\mathbf{U}_{\mathbf{w}}] \implies f = \sigma(f) \mathcal{M}[\mathbf{U}_{\mathbf{w}}], \quad g = \sigma(f) \mathcal{N}[\mathbf{U}_{\mathbf{w}}], \quad h = \sigma(f) \mathcal{P}[\mathbf{U}_{\mathbf{w}}].$$

✎ NB : dans un cas plus général, le flux de masse à travers la paroi s'écrit

$$\int_{\mathbb{R}^3} \left( (\mathbf{v} \cdot \mathbf{n})^+ F_{|\Gamma}^{n,-} + (\mathbf{v} \cdot \mathbf{n})^- F_{|\Gamma}^{n,+} \right) dv_x dv_y dv_z$$

avec  $F_{|\Gamma}^{n,-}$  et  $F_{|\Gamma}^{n,+}$  les valeurs de  $F$  à gauche et à droite de l'interface et  $\mathbf{n}$  la normale à la paroi ; les vitesses  $v_y$  et  $v_z$  apparaissent alors dans l'intégrale : les distributions marginales  $g$  et  $h$  sont donc impactées, de façon non triviale.  $\square$

#### 4.1.5 Système final d'équations

Le modèle se compose de 3 équations décrivant l'évolution des 3 distributions réduites  $f$ ,  $g$  et  $h$  :

$$\begin{cases} \frac{\partial f}{\partial t} + v_x \frac{\partial f}{\partial x} = \frac{1}{\tau} (\mathcal{M}[\mathbf{U}] - f), \\ \frac{\partial g}{\partial t} + v_x \frac{\partial g}{\partial x} = \frac{1}{\tau} (\mathcal{N}[\mathbf{U}] - g), \\ \frac{\partial h}{\partial t} + v_x \frac{\partial h}{\partial x} = \frac{1}{\tau} (\mathcal{P}[\mathbf{U}] - h), \end{cases} \quad (4.4)$$

avec les équilibres associés à chacune de ces distributions :

$$\begin{cases} \mathcal{M}[\mathbf{U}] = \int_{\mathbb{R}^2} \mathcal{M}[\mathbf{U}] dv_y dv_z = \frac{\rho}{\sqrt{2\pi RT}} \exp\left(-\frac{|u - v_x|^2}{2RT}\right), \\ \mathcal{N}[\mathbf{U}] = \left(RT + \frac{v^2}{2}\right) \mathcal{M}[\mathbf{U}], \\ \mathcal{P}[\mathbf{U}] = v \mathcal{M}[\mathbf{U}], \end{cases} \quad (4.5)$$

et les conditions aux limites

$$\begin{cases} f_{|\Gamma} = \sigma(f) \mathcal{M}[\mathbf{U}_{\mathbf{w}}], \quad \sigma(f) = - \left( \int_{\mathbb{R}} v_x^- f dv_x \right) \left( \int_{\mathbb{R}} v_x^+ \mathcal{M}[\mathbf{U}_{\mathbf{w}}] dv_x \right)^{-1}, \\ g_{|\Gamma} = \sigma(f) \mathcal{N}[\mathbf{U}_{\mathbf{w}}], \\ h_{|\Gamma} = \sigma(f) \mathcal{P}[\mathbf{U}_{\mathbf{w}}]; \end{cases} \quad (4.6)$$

les quantités conservatives s'écrivent dans ce nouveau modèle

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho e \end{pmatrix} = \int_{\mathbb{R}} \tilde{\mathbf{m}}(\mathbf{v}) f dv_x + \left( \int_{\mathbb{R}} g dv_x \right) \mathbf{e}_4 + \left( \int_{\mathbb{R}} h dv_x \right) \mathbf{e}_3, \quad \tilde{\mathbf{m}}(\mathbf{v}) = \begin{pmatrix} 1 \\ v_x \\ 0 \\ \frac{1}{2} v_x^2 \end{pmatrix}. \quad (4.7)$$

avec

$$\mathbf{e}_3 = (0, 0, 1, 0)^T, \quad \mathbf{e}_4 = (0, 0, 0, 1)^T.$$

## 4.2 Schéma implicite linéarisé

On décrit ici le schéma numérique implicite appliqué au modèle qui vient d'être exposé. Ce schéma à vitesses discrètes, analogue à celui proposé par Luc Mieussens dans [54], pourra être utilisé avec des méthodes Volumes Finis ou Galerkin Discontinu : la différence entre ces deux méthodes se situe au niveau des termes de transport, les termes de relaxation étant locaux en espace. On écrit donc ici le schéma sous forme semi-discrète en espace, la discrétisation des termes de transport étant abordée par la suite pour chaque type de schéma.

### 4.2.1 Écriture et linéarisation

On écrit le système (4.4) pour le modèle à vitesses discrètes :

$$\forall k \in \mathcal{K}, \quad \begin{cases} \frac{\partial f_k}{\partial t} + v_k \frac{\partial f_k}{\partial x} = \frac{1}{\tau} (\mathcal{M}_k[\mathbf{U}] - f_k), \\ \frac{\partial g_k}{\partial t} + v_k \frac{\partial g_k}{\partial x} = \frac{1}{\tau} (\mathcal{N}_k[\mathbf{U}] - g_k), \\ \frac{\partial h_k}{\partial t} + v_k \frac{\partial h_k}{\partial x} = \frac{1}{\tau} (\mathcal{P}_k[\mathbf{U}] - h_k), \end{cases} \quad (4.8)$$

avec  $f_k(t, \mathbf{x}) = f(t, \mathbf{x}, v_k)$ , et de même pour  $g_k$  et  $h_k$ .

On utilise pour la discrétisation en temps un schéma d'Euler implicite, inconditionnellement stable aussi bien pour une méthode Volumes Finis que pour un schéma Galerkin Discontinu. L'implicitation du temps de relaxation  $\tau$  étant complexe et jugée non critique pour la convergence des calculs, ce terme est exprimé de façon explicite. Le schéma s'écrit donc :

$$\begin{cases} \frac{f_{i,k}^{n+1} - f_{i,k}^n}{\Delta t} + v_k \frac{\partial f_{i,k}^{n+1}}{\partial x} = \frac{1}{\tau_i^n} (\mathcal{M}_k[\mathbf{U}_i^{n+1}] - f_{i,k}^{n+1}), \\ \frac{g_{i,k}^{n+1} - g_{i,k}^n}{\Delta t} + v_k \frac{\partial g_{i,k}^{n+1}}{\partial x} = \frac{1}{\tau_i^n} (\mathcal{N}_k[\mathbf{U}_i^{n+1}] - g_{i,k}^{n+1}), \\ \frac{h_{i,k}^{n+1} - h_{i,k}^n}{\Delta t} + v_k \frac{\partial h_{i,k}^{n+1}}{\partial x} = \frac{1}{\tau_i^n} (\mathcal{P}_k[\mathbf{U}_i^{n+1}] - h_{i,k}^{n+1}). \end{cases}$$

On linéarise l'équilibre autour des moments de la maxwellienne explicite :

$$\mathcal{M}_k[\mathbf{U}_i^{n+1}] = \mathcal{M}_k[\mathbf{U}_i^n] + d\mathcal{M}_k[\mathbf{U}_i^n] + \mathcal{O}((\delta \mathbf{U}_i^n)^2), \quad \delta \mathbf{U}_i^n = \mathbf{U}_i^{n+1} - \mathbf{U}_i^n,$$

et de même pour  $\mathcal{N}_k[\mathbf{U}_i^{n+1}]$  et  $\mathcal{P}_k[\mathbf{U}_i^{n+1}]$  :

$$\begin{cases} \mathcal{N}_k[\mathbf{U}_i^{n+1}] = \mathcal{N}_k[\mathbf{U}_i^n] + d\mathcal{N}_k[\mathbf{U}_i^n] + \mathcal{O}(\delta \mathbf{U}^2), \\ \mathcal{P}_k[\mathbf{U}_i^{n+1}] = \mathcal{P}_k[\mathbf{U}_i^n] + d\mathcal{P}_k[\mathbf{U}_i^n] + \mathcal{O}(\delta \mathbf{U}^2). \end{cases}$$

En posant  $\delta f_{i,k}^n = f_{i,k}^{n+1} - f_{i,k}^n$  (et idem sur  $g$  et  $h$ ), on écrit le schéma sous la " $\delta$ -forme" :

$$\begin{cases} \frac{\delta f_{i,k}^n}{\Delta t} + v_k \frac{\partial \delta f_{i,k}^n}{\partial x} - \frac{1}{\tau_i^n} (\mathcal{M}_k[\mathbf{U}_i^n] - \delta f_{i,k}^n) = -v_k \frac{\partial f_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{M}_k[\mathbf{U}_i^n] - f_{i,k}^n), \\ \frac{\delta g_{i,k}^n}{\Delta t} + v_k \frac{\partial \delta g_{i,k}^n}{\partial x} - \frac{1}{\tau_i^n} (\mathcal{N}_k[\mathbf{U}_i^n] - \delta g_{i,k}^n) = -v_k \frac{\partial g_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{N}_k[\mathbf{U}_i^n] - g_{i,k}^n), \\ \frac{\delta h_{i,k}^n}{\Delta t} + v_k \frac{\partial \delta h_{i,k}^n}{\partial x} - \frac{1}{\tau_i^n} (\mathcal{P}_k[\mathbf{U}_i^n] - \delta h_{i,k}^n) = -v_k \frac{\partial h_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{P}_k[\mathbf{U}_i^n] - h_{i,k}^n). \end{cases}$$

On peut alors écrire le schéma sous la forme du système linéaire  $\left(\frac{\bar{\bar{I}}}{\Delta t} + \bar{\bar{T}} + \bar{\bar{R}}^n\right) \delta \mathbf{F}^n = \mathbf{RHS}^n$ , où les matrices de transport  $\bar{\bar{T}}$  et de relaxation  $\bar{\bar{R}}^n$  et le vecteur second membre  $\mathbf{RHS}^n$  sont définis par :

$$\begin{cases} \delta \mathbf{F}^n = (\delta f^n, \delta g^n, \delta h^n), \\ (\bar{\bar{T}} \delta \mathbf{F}^n)_{i,k} = \left( v_k \frac{\partial \delta f_{i,k}^n}{\partial x}, v_k \frac{\partial \delta g_{i,k}^n}{\partial x}, v_k \frac{\partial \delta h_{i,k}^n}{\partial x} \right), \\ (\bar{\bar{R}}^n \delta \mathbf{F}^n)_{i,k} = \left( -\frac{1}{\tau_i^n} [\mathcal{M}_k[\mathbf{U}_i^n] - \delta f_{i,k}^n], -\frac{1}{\tau_i^n} [\mathcal{N}_k[\mathbf{U}_i^n] - \delta g_{i,k}^n], -\frac{1}{\tau_i^n} [\mathcal{P}_k[\mathbf{U}_i^n] - \delta h_{i,k}^n] \right), \\ (\mathbf{RHS}^n)_{i,k} = \left( -v_k \frac{\partial f_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{M}_k[\mathbf{U}_i^n] - f_{i,k}^n), -v_k \frac{\partial g_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{N}_k[\mathbf{U}_i^n] - g_{i,k}^n), \right. \\ \quad \left. -v_k \frac{\partial h_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{P}_k[\mathbf{U}_i^n] - h_{i,k}^n) \right). \end{cases} \quad (4.9)$$

La présence des termes d'équilibre dans la matrice de relaxation implique un couplage entre les vitesses. Afin de les découpler, et ainsi de résoudre le système vitesse par vitesse, on sépare les termes diagonaux et extradiagonaux de la matrice  $\bar{\bar{R}}^n$ , ces derniers étant envoyés au second membre (méthode de Jacobi) :

$$\left(\frac{\bar{\bar{I}}}{\Delta t} + \bar{\bar{T}} + \bar{\bar{\Delta}}^n\right) \delta \mathbf{F}^n = \mathbf{RHS}^n - \bar{\bar{E}}^n \delta \mathbf{F}^n. \quad (4.10)$$

L'expression des termes extradiagonaux regroupés dans la matrice  $\bar{\bar{E}}^n$  est alors :

$$\begin{aligned} (-\bar{\bar{E}}^n \delta \mathbf{F}^n)_{i,k} = & \left( \frac{1}{\tau_i^n} [\mathcal{M}_k[\mathbf{U}_i^n] - \delta f_{i,k}^n] + \bar{\bar{\Delta}}_{i,k}^{n,f} \delta f_{i,k}^n, \frac{1}{\tau_i^n} [\mathcal{N}_k[\mathbf{U}_i^n] - \delta g_{i,k}^n] + \bar{\bar{\Delta}}_{i,k}^{n,g} \delta g_{i,k}^n, \right. \\ & \left. \frac{1}{\tau_i^n} [\mathcal{P}_k[\mathbf{U}_i^n] - \delta h_{i,k}^n] + \bar{\bar{\Delta}}_{i,k}^{n,h} \delta h_{i,k}^n \right). \end{aligned}$$

La structure du schéma étant posée, il reste à définir l'expression des termes de relaxation apparaissant dans les matrices  $\bar{\bar{\Delta}}^n$  et  $\bar{\bar{E}}^n$  ainsi que dans le second membre  $\mathbf{RHS}^n$ .

## 4.2.2 Expression des différents termes du schéma

### 4.2.2.1 Distributions d'équilibre discrètes

Les trois équilibres discrets  $\mathcal{M}_k$ ,  $\mathcal{N}_k$  et  $\mathcal{P}_k$  sont calculés de façon à préserver les lois de conservation et la propriété d'entropie au niveau discret, comme exposé dans [55].

On écrit  $\mathcal{M}_k$  avec le vecteur des invariants collisionnels  $\mathbf{m}_k$  :

$$\mathcal{M}_k[\mathbf{U}_i^n] = \exp(\boldsymbol{\alpha}_i^n \cdot \mathbf{m}_k), \quad \mathbf{m}_k = (1, v_k, \frac{1}{2}|v_k|^2)^T,$$

où  $\boldsymbol{\alpha}_i^n$  est la variable entropique, qui correspond pour un équilibre continu à

$$\boldsymbol{\alpha} = \left( \ln \left( \frac{\rho}{\sqrt{2\pi RT}} \right) - \frac{u^2}{2RT}, \frac{u}{RT}, \frac{-1}{RT} \right)^T. \quad (4.11)$$

On obtient donc selon les relations (4.5) :

$$\begin{cases} \mathcal{M}_k[\mathbf{U}_i^n] = \exp(\boldsymbol{\alpha}_i^n \cdot \mathbf{m}_k), \\ \mathcal{N}_k[\mathbf{U}_i^n] = \left( \frac{-1}{\alpha_{i,3}^n} + \frac{(v_i^n)^2}{2} \right) \mathcal{M}_k[\mathbf{U}_i^n], \\ \mathcal{P}_k[\mathbf{U}_i^n] = v_i^n \mathcal{M}_k[\mathbf{U}_i^n]. \end{cases} \quad (4.12)$$

Le vecteur des quantités conservatives (4.7) s'écrit au niveau discret :

$$\mathbf{U}(t^n, \mathbf{x}_i) \approx \mathbf{U}_i^n = \sum_k (\tilde{\mathbf{m}}_k f_{i,k}^n + g_{i,k}^n \mathbf{e}_4 + h_{i,k}^n \mathbf{e}_3) \Delta v_k, \quad \text{avec } \tilde{\mathbf{m}}_k = (1, v_k, 0, \frac{1}{2}|v_k|^2)^T$$

On cherche alors à construire les équilibres discrets  $\mathcal{M}_k, \mathcal{N}_k$  et  $\mathcal{P}_k$  de façon à vérifier les lois de conservation au niveau discret, soit

$$\sum_k (\tilde{\mathbf{m}}_k f_{i,k}^n + g_{i,k}^n \mathbf{e}_4 + h_{i,k}^n \mathbf{e}_3) \Delta v_k = \sum_k (\tilde{\mathbf{m}}_k \mathcal{M}_k[\mathbf{U}_i^n] + \mathcal{N}_k[\mathbf{U}_i^n] \mathbf{e}_4 + \mathcal{P}_k[\mathbf{U}_i^n] \mathbf{e}_3) \Delta v_k ;$$

avec les relations (4.12), on obtient ainsi le système non linéaire vérifié par la variable entropique  $\boldsymbol{\alpha}_i^n$  :

$$\sum_k \left[ \tilde{\mathbf{m}}_k \exp(\boldsymbol{\alpha}_i^n \cdot \mathbf{m}_k) + \left( \frac{-1}{\alpha_{i,3}^n} + \frac{(v_i^n)^2}{2} \right) \exp(\boldsymbol{\alpha}_i^n \cdot \mathbf{m}_k) \mathbf{e}_4 + v_i^n \exp(\boldsymbol{\alpha}_i^n \cdot \mathbf{m}_k) \mathbf{e}_3 \right] \Delta v_k = \mathbf{U}_i^n,$$

que l'on peut écrire :

$$\sum_k \tilde{\mathbf{m}}_k \exp(\boldsymbol{\alpha}_i^n \cdot \mathbf{m}_k) \Delta v_k + \left( \frac{-1}{\alpha_{i,3}^n} + \frac{(v_i^n)^2}{2} \right) \rho_i^n \mathbf{e}_4 + \rho_i^n v_i^n \mathbf{e}_3 = \mathbf{U}_i^n.$$

La 3<sup>e</sup> équation de ce système est dégénérée (elle s'écrit  $\rho_i^n v_i^n = \rho_i^n v_i^n$ ) : on le réécrit alors :

$$\sum_k \mathbf{m}_k \exp(\boldsymbol{\alpha}_i^n \cdot \mathbf{m}_k) \Delta v_k - \frac{1}{\alpha_{i,3}^n} \rho_i^n \mathbf{e}_3 = \tilde{\mathbf{U}}_i^n = \left( \rho_i^n, \rho_i^n u_i^n, \rho_i^n \left( e_i^n - \frac{(v_i^n)^2}{2} \right) \right)^T.$$

On résout alors ce système par un algorithme de Newton : en posant

$$\mathbf{F}(\alpha_i^n) = \sum_k \mathbf{m}_k \exp(\alpha_i^n \cdot \mathbf{m}_k) \Delta v_k - \frac{1}{\alpha_{i,3}^n} \rho \mathbf{e}_3 - \tilde{\mathbf{U}}_i^n,$$

on approche  $\alpha_i^n$  par la suite  $(\alpha_{i,p}^n)_p$  définie par :

$$\mathbf{F}'(\alpha_{i,p}^n)(\alpha_{i,p+1}^n - \alpha_{i,p}^n) = -\mathbf{F}(\alpha_{i,p}^n),$$

où  $\alpha_{i,0}^n$  est donné par (4.11) et où la jacobienne  $\mathbf{F}'(\alpha_{i,p}^n)$  s'écrit :

$$\mathbf{F}'(\alpha_{i,p}^n) = \frac{\partial \mathbf{F}}{\partial \alpha}(\alpha_{i,p}^n) = \sum_k \left[ \frac{\partial}{\partial \alpha} \left( \mathbf{m}_k \exp(\alpha \cdot \mathbf{m}_k) \right) (\alpha_{i,p}^n) - \frac{\partial}{\partial \alpha} \left( \frac{\rho}{\alpha_3} \mathbf{e}_3 \right) (\alpha_{i,p}^n) \right] \Delta v_k.$$

Or :

$$\begin{aligned} \sum_k \frac{\partial}{\partial \alpha} \left( \mathbf{m}_k \exp(\alpha \cdot \mathbf{m}_k) \right) \Delta v_k &= \sum_k \left[ \mathbf{m}_{kj} \frac{\partial}{\partial \alpha_i} (\exp(\alpha \cdot \mathbf{m}_k)) \right]_{i,j} \Delta v_k \\ &= \sum_k \left( \frac{\partial}{\partial \alpha} (\exp(\alpha \cdot \mathbf{m}_k)) \otimes \mathbf{m}_k \right) \Delta v_k \\ &= \sum_k \left( \mathbf{m}_k \otimes \mathbf{m}_k \exp(\alpha \cdot \mathbf{m}_k) \right) \Delta v_k, \end{aligned}$$

et :

$$\frac{\partial}{\partial \alpha} \left( \frac{\rho}{\alpha_3} \mathbf{e}_3 \right) = \left[ \frac{\partial}{\partial \alpha_i} \left( \left( \frac{\rho}{\alpha_3} \right) \delta_{j,3} \right) \right]_{i,j} = \left[ -\frac{\rho}{\alpha_3^2} \delta_{i,3} \delta_{j,3} \right]_{i,j} = -\frac{\rho}{\alpha_3^2} \mathbf{e}_3 \otimes \mathbf{e}_3,$$

ce qui donne finalement :

$$\mathbf{F}'(\alpha_{i,p}^n) = \sum_k \left( \mathbf{m}_k \otimes \mathbf{m}_k \exp(\alpha_{i,p}^n \cdot \mathbf{m}_k) \right) \Delta v_k + \frac{\rho}{(\alpha_{i,p,3}^n)^2} \mathbf{e}_3 \otimes \mathbf{e}_3$$

#### 4.2.2.2 Différentielles des distributions d'équilibre

La manière la plus naturelle d'exprimer la différentielle des différents équilibres est d'utiliser le vecteur des variables conservatives  $\mathbf{U}$ , mais le calcul n'est pas trivial. On passe alors par les variables thermodynamiques  $\rho, u, v$  et  $e$  :

$$d\mathcal{M}_k[\mathbf{U}_i^n] = \frac{\partial \mathcal{M}_k}{\partial \rho}[\mathbf{U}_i^n] \cdot \delta \rho_i^n + \frac{\partial \mathcal{M}_k}{\partial u}[\mathbf{U}_i^n] \cdot \delta u_i^n + \frac{\partial \mathcal{M}_k}{\partial v}[\mathbf{U}_i^n] \cdot \delta v_i^n + \frac{\partial \mathcal{M}_k}{\partial e}[\mathbf{U}_i^n] \cdot \delta e_i^n. \quad (4.13)$$

Ces 4 variables sont liées par la température, selon la relation (4.3), ce qui donne :

$$\frac{\partial T}{\partial \rho} = 0, \quad \frac{\partial T}{\partial u} = \frac{-2u}{3R}, \quad \frac{\partial T}{\partial v} = \frac{-2v}{3R}, \quad \frac{\partial T}{\partial e} = \frac{2}{3R}.$$

Pour simplifier encore les calculs, on approche les dérivées des équilibres discrets par celles des équilibres continus, suivant une idée de Titarev [68] : on suppose

$$d\mathcal{M}_k[\mathbf{U}_i^n] \approx d\mathcal{M}[\mathbf{U}_i^n, v_k].$$



On a alors :

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{M}}{\partial T} = \frac{1}{T} \left( \frac{(u - v_x)^2}{2RT} - \frac{1}{2} \right) \mathcal{M}, \\ \frac{\partial \mathcal{M}}{\partial \rho} = \frac{1}{\rho} \mathcal{M}, \\ \frac{\partial \mathcal{M}}{\partial u} = \left( \frac{v_x - u}{RT} \right) \mathcal{M} + \frac{\partial \mathcal{M}}{\partial T} \frac{\partial T}{\partial u} = \left( \frac{v_x - u}{RT} \right) \mathcal{M} - \frac{2u}{3R} \frac{\partial \mathcal{M}}{\partial T}, \\ \frac{\partial \mathcal{M}}{\partial v} = \frac{\partial \mathcal{M}}{\partial T} \frac{\partial T}{\partial v} = -\frac{2v}{3R} \frac{\partial \mathcal{M}}{\partial T}, \\ \frac{\partial \mathcal{M}}{\partial e} = \frac{\partial \mathcal{M}}{\partial T} \frac{\partial T}{\partial e} = \frac{2}{3R} \frac{\partial \mathcal{M}}{\partial T}, \end{array} \right. \quad (4.14)$$

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{N}}{\partial \rho} = \left( RT + \frac{v^2}{2} \right) \frac{\partial \mathcal{M}}{\partial \rho} = \frac{1}{\rho} \left( RT + \frac{v^2}{2} \right) \mathcal{M}, \\ \frac{\partial \mathcal{N}}{\partial u} = \left( RT + \frac{v^2}{2} \right) \frac{\partial \mathcal{M}}{\partial u} + R \frac{\partial T}{\partial u} \mathcal{M} = \left( RT + \frac{v^2}{2} \right) \frac{\partial \mathcal{M}}{\partial u} - \frac{2u}{3} \mathcal{M}, \\ \frac{\partial \mathcal{N}}{\partial v} = \left( RT + \frac{v^2}{2} \right) \frac{\partial \mathcal{M}}{\partial v} + \left( R \frac{\partial T}{\partial v} + v \right) \mathcal{M} = \left( RT + \frac{v^2}{2} \right) \frac{\partial \mathcal{M}}{\partial v} + \frac{v}{3} \mathcal{M}, \\ \frac{\partial \mathcal{N}}{\partial e} = \left( RT + \frac{v^2}{2} \right) \frac{\partial \mathcal{M}}{\partial e} + R \frac{\partial T}{\partial e} \mathcal{M} = \left( RT + \frac{v^2}{2} \right) \frac{\partial \mathcal{M}}{\partial e} + \frac{2}{3} \mathcal{M}, \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{P}}{\partial \rho} = v \frac{\partial \mathcal{M}}{\partial \rho} = \frac{v}{\rho} \mathcal{M}, \\ \frac{\partial \mathcal{P}}{\partial u} = v \frac{\partial \mathcal{M}}{\partial u}, \\ \frac{\partial \mathcal{P}}{\partial v} = v \frac{\partial \mathcal{M}}{\partial v} + \mathcal{M}, \\ \frac{\partial \mathcal{P}}{\partial e} = v \frac{\partial \mathcal{M}}{\partial e}. \end{array} \right.$$

Les incréments  $\delta \rho_i^n, \delta u_i^n, \delta v_i^n, \delta e_i^n$  sont calculés de la façon suivante :

$$\left\{ \begin{array}{l} \delta \rho_i^n = \sum_{k'} \delta f_{i,k'}^n \Delta v_{k'}, \\ \delta u_i^n = \frac{1}{\rho_i^{n+1}} \left( \delta(\rho u)_i^n - u_i^n \delta \rho_i^n \right) = \frac{1}{\rho_i^n + \delta \rho_i^n} \left( \delta(\rho u)_i^n - u_i^n \delta \rho_i^n \right) \\ \quad = \frac{1}{\rho_i^n} \left( \delta(\rho u)_i^n - u_i^n \delta \rho_i^n \right) + \mathcal{O} \left( \delta(\rho_i^n)^2 \right) \approx \frac{1}{\rho_i^n} \left( \sum_{k'} v_{k'} \delta f_{i,k'}^n \Delta v_{k'} - u_i^n \delta \rho_i^n \right), \\ \delta v_i^n = \frac{1}{\rho_i^{n+1}} \left( \delta(\rho v)_i^n - v_i^n \delta \rho_i^n \right) \approx \frac{1}{\rho_i^n} \left( \sum_{k'} \delta h_{i,k'}^n \Delta v_{k'} - v_i^n \delta \rho_i^n \right), \\ \delta e_i^n = \frac{1}{\rho_i^{n+1}} \left( \delta(\rho e)_i^n - e_i^n \delta \rho_i^n \right) \approx \frac{1}{\rho_i^n} \left( \sum_{k'} \left( \frac{1}{2} |v_{k'}|^2 \delta f_{i,k'}^n + \delta g_{i,k'}^n \right) \Delta v_{k'} - e_i^n \delta \rho_i^n \right). \end{array} \right. \quad (4.15)$$

### 4.2.2.3 Termes diagonaux de la matrice de relaxation $\bar{\bar{R}}^n$ (matrice $\bar{\bar{\Delta}}^n$ )

D'après (4.14), (4.13) et (4.15), on peut réécrire la première composante du terme  $(\bar{\bar{R}}^n \delta \mathbf{F}^n)_{i,k}$  apparaissant dans (4.9)

$$\begin{aligned} (\bar{\bar{R}}^n \delta f)_{i,k} &= -\frac{1}{\tau_i^n} (\mathcal{M}_k - \delta f_{i,k}^n) = -\frac{1}{\tau_i^n} \left[ \frac{\partial \mathcal{M}_k}{\partial \rho} [\mathbf{U}_i^n] \cdot \sum_l \delta f_{i,l}^n \Delta v_l \right. \\ &\quad + \frac{\partial \mathcal{M}_k}{\partial u} [\mathbf{U}_i^n] \cdot \frac{1}{\rho_i^n} \left( \sum_l (v_l - u_i^n) \delta f_{i,l}^n \Delta v_l \right) + \frac{\partial \mathcal{M}_k}{\partial v} [\mathbf{U}_i^n] \cdot \frac{1}{\rho_i^n} \left( \sum_l (\delta h_{i,l}^n - v_i^n \delta f_{i,l}^n) \Delta v_l \right) \\ &\quad \left. + \frac{\partial \mathcal{M}_k}{\partial e} [\mathbf{U}_i^n] \cdot \frac{1}{\rho_i^n} \left( \sum_l \left( \frac{1}{2} |v_l|^2 - e_i^n \right) \delta f_{i,l}^n + \delta g_{i,l}^n \right) \Delta v_l \right] - \delta f_{i,k}^n. \end{aligned}$$

La contribution du terme  $\delta f_{i,k}^n$  est donc  $\Delta_{i,k}^{n,f} \delta f_{i,k}^n$ , avec

$$\begin{aligned} \Delta_{i,k}^{n,f} &= -\frac{1}{\tau_i^n} \left[ \left( \frac{\partial \mathcal{M}_k}{\partial \rho} [\mathbf{U}_i^n] + \frac{1}{\rho_i^n} \left( (v_k - u_i^n) \frac{\partial \mathcal{M}_k}{\partial u} [\mathbf{U}_i^n] - v_i^n \frac{\partial \mathcal{M}_k}{\partial v} [\mathbf{U}_i^n] \right. \right. \right. \\ &\quad \left. \left. \left. + \left( \frac{1}{2} |v_k|^2 - e_i^n \right) \frac{\partial \mathcal{M}_k}{\partial e} [\mathbf{U}_i^n] \right) \right) \Delta v_k - 1 \right]. \end{aligned}$$

L'expression des termes  $(\bar{\bar{R}}^n \delta g)_{i,k}$  et  $(\bar{\bar{R}}^n \delta h)_{i,k}$  est analogue mais plus simple, ce qui donne :

$$\begin{cases} \Delta_{i,k}^{n,g} = -\frac{1}{\tau_i^n} \left( \frac{1}{\rho_i^n} \frac{\partial \mathcal{N}_k}{\partial e} [\mathbf{U}_i^n] \Delta v_k - 1 \right), \\ \Delta_{i,k}^{n,h} = -\frac{1}{\tau_i^n} \left( \frac{1}{\rho_i^n} \frac{\partial \mathcal{P}_k}{\partial u} [\mathbf{U}_i^n] \Delta v_k - 1 \right). \end{cases}$$

Il reste à expliciter les termes de transport, différents pour chaque type de schéma. On les détaille pour les schémas Volumes Finis dans la partie 4.3, puis pour le schéma Galerkin Discontinu dans la partie 4.4.

## 4.3 Schémas Volumes Finis : flux et condition de réflexion diffuse

Les termes advectifs font intervenir les conditions aux limites imposées en bordure du domaine de calcul : la discrétisation de ces deux éléments se traite ensemble.

On prend numériquement en compte ces conditions aux limites par une stratégie de mailles fictives. Selon le stencil du schéma considéré, on introduit une ou plusieurs mailles fictives pour chaque bord, et les valeurs correspondantes des fonctions de distribution. On détaille dans cette partie la construction de ces conditions aux limites numériques.

### 4.3.1 Schémas d'ordre 1

Le terme de transport de  $f$  s'intègre dans la maille  $C_i$  :

$$\frac{1}{\Delta x_i} \int_{C_i} v_k \frac{\partial f_{i,k}^n}{\partial x} dx = \frac{1}{\Delta x_i} \left( \mathcal{F}_{i+\frac{1}{2},k}^n - \mathcal{F}_{i-\frac{1}{2},k}^n \right), \quad (4.16)$$

où  $\mathcal{F}_{i+\frac{1}{2},k}^n$  est le flux numérique à travers l'interface en  $x_{i+\frac{1}{2}}$ . On choisit un flux upwind :

$$\mathcal{F}_{i+\frac{1}{2},k}^n = v_k^+ f_{i,k}^n + v_k^- f_{i+1,k}^n,$$

avec la notation classique des parties positive et négative :

$$v_k^+ = \frac{1}{2} (v_k + |v_k|), \quad v_k^- = \frac{1}{2} (v_k - |v_k|).$$

Le schéma ayant un stencil à 3 points, on n'utilise qu'une seule maille fictive à chaque bord. Considérons la maille fictive en  $i=0$  ; la condition d'équilibre thermodynamique à la paroi décrite en (4.6) s'obtient en posant :

$$\forall k \mid v_k > 0, \quad f_{0,k}^n = \sigma_f \mathcal{M}_k^n, \quad (4.17)$$

où  $\sigma_f$  est une constante calculée de façon à vérifier la condition sur le flux de masse :

$$\sigma_f = - \left( \sum_l v_l^- f_{1,l}^n \Delta v_l \right) \left( \sum_l v_l^+ \mathcal{M}_l^n \Delta v_l \right)^{-1}$$

ce qui donne bien :  $\mathcal{F}_{1/2}^n = \sum_k \mathcal{F}_{1/2,k}^n \Delta v_k = \sum_k \left( -v_k^+ \frac{\sum_l v_l^- f_{1,l}^n \Delta v_l}{\sum_l v_l^+ \mathcal{M}_l^n \Delta v_l} \mathcal{M}_k^n + v_k^- f_{1,k}^n \right) \Delta v_k = 0$ .

On vérifie ainsi simultanément la condition de flux de masse nul et la condition d'équilibre thermodynamique.

Pour  $v_k < 0$ , le sens physique donne  $f_{0,k}^n = f_{1,k}^n$ . Cependant, la valeur de  $f_{0,k}^n$  pour  $v_k < 0$  n'a pas d'importance, cet élément n'intervenant pas dans le schéma.

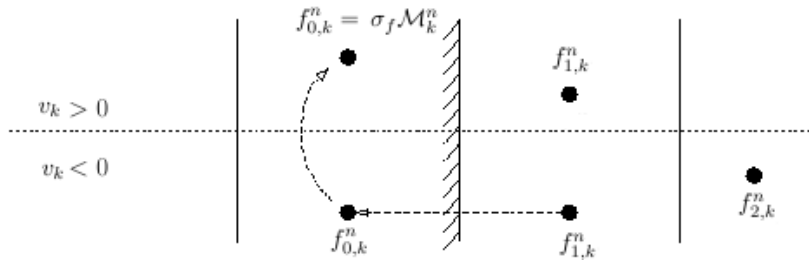


FIGURE 4.2: Construction des cellules fictives pour  $f$  à l'ordre 1

Conformément à la condition imposée aux distributions  $g$  et  $h$ , on pose :

$$\begin{cases} g_{0,k}^n = \sigma_f \mathcal{N}_k^n & \text{et } h_{0,k}^n = \sigma_f \mathcal{P}_k^n & \text{si } v_k > 0, \\ g_{0,k}^n = g_{1,k}^n & \text{et } h_{0,k}^n = h_{1,k}^n & \text{si } v_k < 0. \end{cases}$$

### 4.3.2 Schémas d'ordre 2

On peut augmenter la précision du schéma Volumes Finis d'ordre 1 présenté ci-dessus de plusieurs manières. On en illustre ici deux : la reconstruction linéaire de l'inconnue sur chaque maille, éventuellement avec un limiteur de pente, et l'introduction d'un limiteur de flux. Dans chaque cas, on détaille la construction correspondante des conditions aux limites numériques.

#### 4.3.2.1 Schémas à reconstruction linéaire

On intègre le terme de flux dans chaque maille selon (4.16) et on choisit un flux upwind comme à l'ordre 1, mais en reconstruisant les valeurs de  $f$  aux interfaces du maillage :

$$\begin{cases} \mathcal{F}_{i+\frac{1}{2},k}^n = v_k^+ f_{i+\frac{1}{2},k}^{n,-} + v_k^- f_{i+\frac{1}{2},k}^{n,+}, \\ f_{i+\frac{1}{2},k}^{n,-} = f_{i,k}^n + \frac{\Delta x_i}{2} \delta_{i,k}^n, \\ f_{i+\frac{1}{2},k}^{n,+} = f_{i+1,k}^n - \frac{\Delta x_{i+1}}{2} \delta_{i+1,k}^n, \end{cases} \quad (4.18)$$

où  $\delta_{i,k}^n(f_{i-1,k}^n, f_{i,k}^n, f_{i+1,k}^n)$  est la pente dans la maille  $i$  de la distribution  $f$  reconstruite :

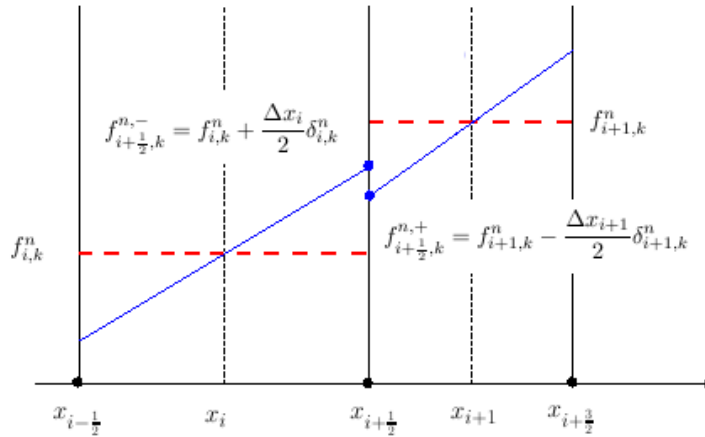


FIGURE 4.3: Reconstruction linéaire de la distribution  $f$

La méthode la plus classique pour calculer cette pente est une approche aux moindres carrés, qui dans le cas d'un maillage de pas constant fournit une expression très simple :

$$\delta_{i,k}^n = \frac{f_{i+1,k}^n - f_{i-1,k}^n}{2\Delta x_i}. \quad (4.19)$$

Le stencil du schéma obtenu compte cette fois 5 points, il faut donc deux cellules fictives à chaque bord.

Pour préserver la monotonie de la reconstruction, on peut utiliser un limiteur. L'expression

générale de la pente limitée est :

$$\delta_{i,k}^{n,lim} = \minmod\left(\delta_{i,k}^n, 2\alpha_i \frac{f_{i,k}^n - f_{i-1,k}^n}{\Delta x_i}, 2\alpha_i \frac{f_{i+1,k}^n - f_{i,k}^n}{\Delta x_i}\right), \quad 0 \leq \alpha_i \leq 1. \quad (4.20)$$

où la fonction  $\minmod \mathbb{R}^3 \rightarrow \mathbb{R}$  est définie par

$$\forall (x, y, z) \in \mathbb{R}^3, \quad \minmod(x, y, z) = \begin{cases} \operatorname{sgn}(x) \min(|x|, |y|, |z|) & \text{si } \operatorname{sgn}(x) = \operatorname{sgn}(y) = \operatorname{sgn}(z), \\ 0 & \text{sinon.} \end{cases}$$

On va supposer dans la suite que  $\Delta x_{-1} = \Delta x_0 = \Delta x_1$ , ce qui donne dans la première cellule fictive :

$$\delta_{0,k}^{n,lim} = \minmod\left(\frac{f_{1,k}^n - f_{-1,k}^n}{2\Delta x_0}, 2\alpha_0 \frac{f_{0,k}^n - f_{-1,k}^n}{\Delta x_0}, 2\alpha_0 \frac{f_{1,k}^n - f_{0,k}^n}{\Delta x_0}\right). \quad (4.21)$$

Si les valeurs  $f_{-1,k}^n, f_{0,k}^n, f_{1,k}^n$  forment une suite non monotone, le limiteur s'annule. Sinon il sélectionne la pente la plus "douce" parmi celles de 3 droites construites avec ces points, comme illustré ci-dessous, les pentes en pointillés étant pondérées du coefficient  $\alpha_0$  :

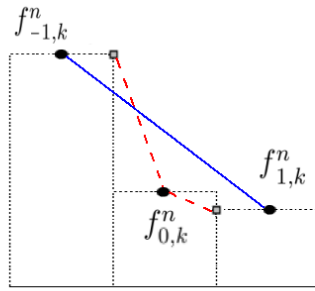


FIGURE 4.4: Limiteur de pente

La construction de la condition aux limites se fait alors suivant la démarche suivante, inspirée de [35] et [30] :

- ÉTAPE 1 : calcul de  $f_{1/2,l}^{n,+}$  pour toutes les vitesses  $v_l < 0$ .

Pour  $v_k < 0$ , la valeur de  $f_{1/2,k}^{n,+}$  est calculée par extrapolation sur le bord du domaine fluide. Le plus naturel est de choisir une extrapolation linéaire à partir de  $f_{1,k}^n$  et  $f_{2,k}^n$ , mais d'autres façons d'extrapoler sont a priori envisageables (notamment pour éviter que la valeur de  $f_{1/2,k}^{n,+}$  ne soit négative). La même extrapolation donne  $f_{0,k}^n$  et  $f_{-1,k}^n$ . On procède de la même façon pour calculer les valeurs de  $g$  et  $h$  pour  $v_k < 0$ .

- ÉTAPE 2 : calcul de  $f_{1/2,k}^{n,-}$  pour toutes les vitesses  $v_l > 0$ .

Comme en (4.17), on annule  $\mathcal{F}_{1/2}^n$  tout en respectant l'équilibre thermodynamique en posant :

$$f_{1/2,k}^{n,-} = -\left(\sum_l v_l^- f_{1/2,l}^{n,+} \Delta v_l\right) \left(\sum_l v_l^+ \mathcal{M}_l^n \Delta v_l\right)^{-1} \mathcal{M}_k^n = \sigma_f \mathcal{M}_k^n. \quad (4.22)$$

On définit  $g_{1/2,k}^{n,-}$  et  $h_{1/2,k}^{n,-}$  de façon analogue à  $f_{1/2,k}^{n,-}$  :

$$\begin{cases} g_{1/2,k}^{n,-} = \sigma_f \mathcal{N}_k^n, \\ h_{1/2,k}^{n,-} = \sigma_f \mathcal{P}_k^n. \end{cases}$$

- ÉTAPE 3 : définition des valeurs  $f_{0,k}^n$  et  $f_{-1,k}^n$  dans les cellules fictives.

Cette construction doit se faire de façon à ce que la valeur de  $f_{1/2,k}^{n,-}$  soit cohérente avec la reconstruction linéaire et la limitation de pente, c'est-à-dire de telle sorte que la condition aux limites soit du même ordre que le schéma. En d'autres termes, pour que la valeur de  $f_{1/2,k}^{n,-}$  définie par (4.22) satisfasse aussi la seconde équation de (4.18), nous cherchons à ce que le limiteur défini en (4.21) vérifie :

$$\delta_{0,k}^{n,lim} = \frac{2}{\Delta x_0} (f_{1/2,k}^{n,-} - f_{0,k}^n). \quad (4.23)$$

La valeur de  $f_{1/2,k}^{n,-}$  étant désormais déterminée, on va construire les points  $f_{-1,k}^n$  et  $f_{0,k}^n$  de façon à forcer le limiteur à respecter cette contrainte. Géométriquement, la figure 4.4 montre que pour imposer une valeur à  $\delta_{0,k}^{n,lim}$ , il suffit d'aligner les points  $f_{-1,k}^n$ ,  $f_{0,k}^n$  et  $f_{1,k}^n$  sur une même droite de pente  $\delta$  : on obtient  $\delta_{0,k}^{n,lim} = \min(\delta, 2\alpha_0\delta, 2\alpha_0\delta)$ . On distingue alors deux cas :

➤ **Cas 1** :  $\alpha_0 \geq 1/2$ . On a  $\delta_{0,k}^{n,lim} = \delta$ .

La condition (4.23) est alors vérifiée si les 4 points  $f_{-1,k}^n$ ,  $f_{0,k}^n$ ,  $f_{1/2,k}^{n,-}$  et  $f_{1,k}^n$  sont alignés, ce qui se traduit par :

$$\begin{cases} f_{0,k}^n = f_{1,k}^n - 2\Delta_f^k, \\ f_{-1,k}^n = f_{1,k}^n - 4\Delta_f^k, \end{cases} \quad (4.24)$$

où l'on note :  $\Delta_f^k = f_{1,k}^n - f_{1/2,k}^{n,-}$ .

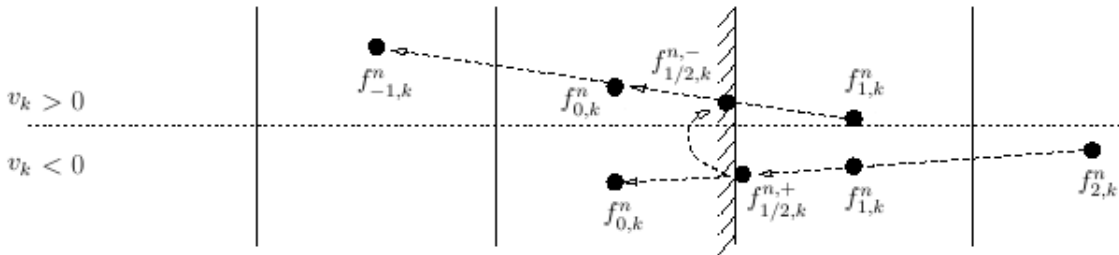


FIGURE 4.5: Construction des cellules fictives à l'ordre 2 pour  $\alpha_0 \geq 1/2$

et suivant le même raisonnement que pour  $f$  :

$$\begin{cases} \Delta_g^k = g_{1,k}^n - g_{1/2,k}^{n,-}, \\ g_{0,k}^n = g_{1,k}^n - 2\Delta_g^k, \\ g_{-1,k}^n = g_{1,k}^n - 4\Delta_g^k \end{cases} \quad \text{et} \quad \begin{cases} \Delta_h^k = h_{1,k}^n - h_{1/2,k}^{n,-}, \\ h_{0,k}^n = h_{1,k}^n - 2\Delta_h^k, \\ h_{-1,k}^n = h_{1,k}^n - 4\Delta_h^k. \end{cases}$$

➤ **Cas 2 :**  $\alpha_0 \leq 1/2$ . On a  $\delta_{0,k}^{n,lim} = 2\alpha_0\delta$ , ce qui donne 3 égalités :

$$\begin{cases} \frac{2}{\Delta x_0} (f_{1/2,k}^{n,-} - f_{0,k}^n) = 2\alpha_0\delta \text{ d'après (4.23),} \\ \frac{1}{\Delta x_0} (f_{1,k}^n - f_{0,k}^n) = \delta \text{ par construction,} \\ \frac{1}{\Delta x_0} (f_{0,k}^n - f_{-1,k}^n) = \delta. \end{cases}$$

Les valeurs  $f_{0,k}^n$  et  $f_{-1,k}^n$  sont donc des moyennes pondérées de  $f_{1/2,k}^{n,-}$  et  $f_{1,k}^n$  :

$$\begin{cases} f_{0,k}^n = \frac{1}{1-\alpha_0} (f_{1/2,k}^{n,-} - \alpha_0 f_{1,k}^n), \\ f_{-1,k}^n = 2f_{0,k}^n - f_{1,k}^n = \frac{1}{1-\alpha_0} (2f_{1/2,k}^{n,-} - (1+\alpha_0)f_{1,k}^n). \end{cases} \quad (4.25)$$

Le point  $f_{1/2,k}^{n,-}$  n'est alors plus aligné avec les trois autres.

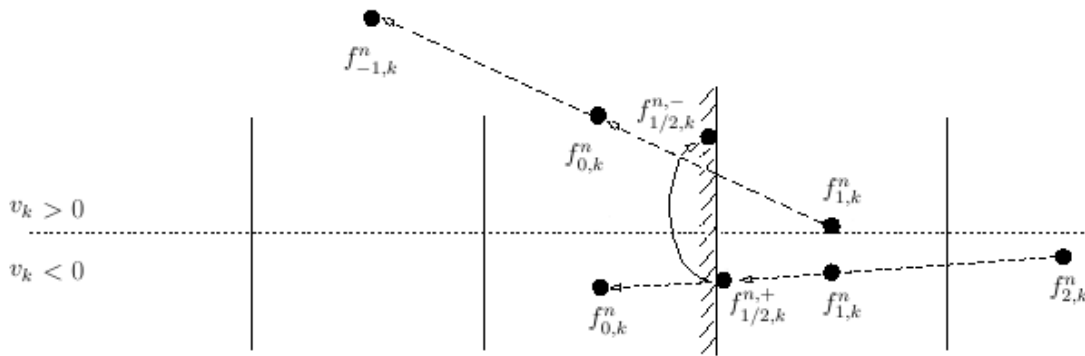


FIGURE 4.6: Construction des cellules fictives à l'ordre 2 pour  $\alpha_0 \leq 1/2$

Comme pour le cas  $\alpha_0 > \frac{1}{2}$ , la définition des valeurs de  $g$  et  $h$  dans les cellules virtuelles est analogue à celle de  $f$ .

Les valeurs de  $f_{-1,k}^n$  et  $f_{0,k}^n$  sont à présent entièrement déterminées par l'étape 1 et par (4.24) ou (4.25).

(NB : il n'y a a priori pas unicité de la façon d'imposer la valeur du limiteur; cependant celle choisie ici est cohérente avec la stratégie de reconstruction linéaire qui sous-tend la définition des limiteurs de pente, de même que l'extrapolation linéaire utilisée dans le cas  $v_k < 0$ ).

Considérons maintenant un autre type de schéma d'ordre 2, les schémas à limiteur de flux.

### 4.3.2.2 Schémas à limiteur de flux

Dans ce type de schémas, on ne procède plus par reconstruction de l'inconnue aux interfaces, mais par l'introduction d'un terme de limitation directement dans le flux numérique. On prend par exemple le limiteur de Yee décrit dans [73], qui s'écrit :

$$\begin{cases} \mathcal{F}_{i+\frac{1}{2},k}^n = v_k^+ f_{i,k}^n + v_k^- f_{i+1,k}^n + |v_k| \frac{\Delta x_i}{2} \Phi_{i+1/2,k}^n, \\ \Phi_{i+1/2,k}^n = \min\text{mod}(\Delta f_{i-1/2,k}^n, \Delta f_{i+1/2,k}^n, \Delta f_{i+3/2,k}^n), \\ \Delta f_{i+1/2,k}^n = \frac{f_{i+1,k}^n - f_{i,k}^n}{\Delta x_i}. \end{cases} \quad (4.26)$$

La condition aux limites "naïve" généralement utilisée sur ce schéma est :

$$f_{-1,k}^n = f_{0,k}^n = - \left( \sum_l v_l^- f_{1,l}^n \Delta v_l \right) \left( \sum_l v_l^+ \mathcal{M}_l^n \Delta v_l \right)^{-1} \mathcal{M}_k^n,$$

ce qui donne  $\Phi_{1/2,k}^n = 0$  : on a donc un flux au bord d'ordre 1. Cette condition aux limites est celle actuellement utilisée dans le code d'aérodynamique raréfiée du CEA-CESTA.

Il est donc indispensable de voir si l'on peut construire  $f_{-1,k}^n$  et  $f_{0,k}^n$  de façon à avoir un flux d'ordre 2. Malheureusement, nous allons montrer ci-dessous que cela n'est pas possible en général.

En réécrivant le flux :

$$\mathcal{F}_{i+\frac{1}{2},k}^n = v_k^+ \left( f_{i,k}^n + \frac{\Delta x_i}{2} \Phi_{i+1/2,k}^n \right) + v_k^- \left( f_{i+1,k}^n - \frac{\Delta x_i}{2} \Phi_{i+1/2,k}^n \right),$$

on voit que pour obtenir un flux de masse nul à la paroi, on doit avoir par analogie avec (4.17) :

$$\forall k \mid v_k > 0, f_{0,k}^n + \frac{\Delta x_0}{2} \Phi_{1/2,k}^n = - \frac{\sum_l v_l^- \left( f_{1,l}^n - \frac{\Delta x_0}{2} \Phi_{1/2,l}^n \right) \Delta v_l}{\sum_l v_l^+ \mathcal{M}_l^n \Delta v_l} \mathcal{M}_k^n = \sigma_f \mathcal{M}_k^n,$$

et on retrouve une condition semblable au cas précédent :

$$\forall k \mid v_k > 0, \Phi_{1/2,k}^n = \frac{2}{\Delta x_0} \left( \sigma_f \mathcal{M}_k^n - f_{0,k}^n \right). \quad (4.27)$$

Nous cherchons donc s'il est possible de construire  $f_{0,k}^n$  de sorte que (4.26) et (4.27) soient satisfaits simultanément. Or ici,  $\Phi_{1/2,k}^n$  dépend de  $f_{-1,k}^n, f_{0,k}^n, f_{1,k}^n$  et  $f_{2,k}^n$  : on a un stencil à 4 points, dont deux sont déjà déterminés,  $f_{1,k}^n$  et  $f_{2,k}^n$  (contrairement au cas du limiteur de pente, qui a un stencil à 3 points dont un seul déterminé) ; le signe de  $\Delta f_{3/2,k}^n$  est donc fixé, et  $\Delta f_{1/2,k}^n$  et  $\Delta f_{-1/2,k}^n$  doivent donc avoir le même signe pour éviter que  $\Phi_{1/2,k}^n = 0$ . On a donc une première condition nécessaire :

$$(f_{1,k}^n - f_{0,k}^n)(f_{2,k}^n - f_{1,k}^n) > 0. \quad (4.28)$$

et nous n'utilisons pas la condition sur le signe de  $\Delta f_{-1/2,k}^n$  ici.



Supposons alors que les valeurs de  $f_{-1,k}^n$  et  $f_{0,k}^n$  sont telles que  $\Phi_{1/2,k}^n \neq 0$ , et l'on a donc  $|\Phi_{1/2,k}^n| = \min(|\Delta f_{-1/2,k}^n|, |\Delta f_{1/2,k}^n|, |\Delta f_{3/2,k}^n|)$ , et en particulier,  $|\Phi_{1/2,k}^n| \leq |\Delta f_{3/2,k}^n|$ , ce qui combiné à (4.27) donne une dernière condition nécessaire

$$2|\sigma_f \mathcal{M}_k^n - f_{0,k}^n| \leq |f_{2,k}^n - f_{1,k}^n|. \quad (4.29)$$

On voit sur la figure 4.7 que selon la valeur de  $\sigma_f \mathcal{M}_k^n$ , la satisfaction simultanée de (4.28) et (4.29) n'est pas toujours possible : la première condition impose de choisir  $f_{0,k}^n$  dans la zone (A) (en pointillés), alors que la seconde le contraint à rester dans la zone (B) (en tirets), zones qui peuvent être disjointes :

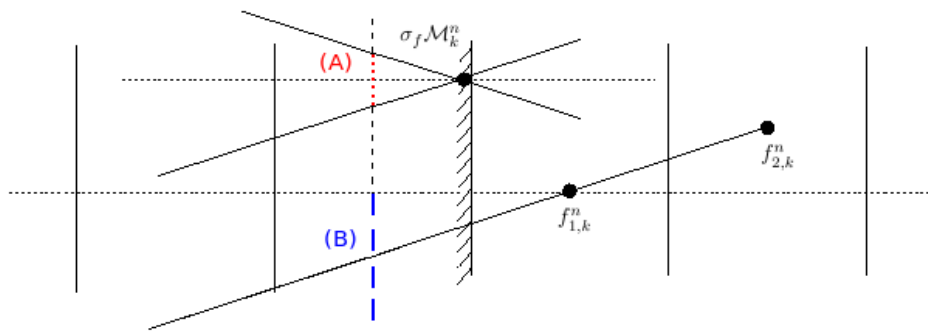


FIGURE 4.7: Cellules fictives pour un schéma à limiteur de flux

Une méthode à limiteur de flux ne peut donc pas donner un flux d'ordre 2 au bord.

#### 4.3.2.3 Comparaison des deux types de schémas

On compare sur l'écoulement de Couette le schéma à limiteur de flux "Yee" (qui est donc d'ordre 1 au bord) avec deux schémas à reconstruction linéaire : le premier effectue une reconstruction sans limiteur "Reconstruction", le second utilise le limiteur avec  $\alpha_i = 1/2$  "Lim-CL2".

**En pratique, la limitation de pente ne fonctionne pas dans le cas  $\alpha_i < 1/2$ , quelle que soit la condition aux limites utilisée :** on observe une stationnarisation du résidu, le schéma ne converge pas, pour une raison inexpliquée. Le cas  $\alpha_i = 1/2$  est fonctionnel avec l'une ou l'autre construction de la condition aux limites, qui dans ce cas sont équivalentes. On doit donc soit se contenter d'une reconstruction linéaire sans limitation au risque de perdre la monotonie, ce qui ne semble pas poser de problème, du moins pour ce cas particulier, soit utiliser un limiteur avec  $\alpha_i \geq 1/2$  (au moins sur les bords du domaine de calcul).

Dans le cas non limité, la pente  $\delta_{0,k}^n$  est donnée par la relation (4.19) et doit vérifier la contrainte (4.23). Il est donc naturel d'utiliser les conditions aux limites décrites dans le cas  $\alpha_0 \geq 1/2$  : les valeurs des fonctions de distribution dans les mailles fictives sont données par les relations (4.24). Le schéma limité utilise les mêmes conditions aux limites.

On montre également le schéma limité utilisé avec des conditions aux limites d'ordre 1 (Lim-CL1), c'est-à-dire avec des valeurs dans les mailles fictives données par les relations (4.17), que l'on adapte à la présence d'une seconde maille fictive en posant  $f_{-1,k}^n = f_{0,k}^n$ .

Les plaques, distantes de  $D = 1$  m, ont une température  $T_w = 273$  K; celle de droite bouge à la vitesse  $v_w = 300$  m.s<sup>-1</sup>, celle de gauche ne bouge pas. Le gaz (argon) est initialement à l'équilibre, immobile et à la même température que les plaques. Le maillage physique est de pas constant et compte 100 cellules, l'espace des vitesses discrètes comprend 40 vitesses uniformément réparties entre  $v_{min}$  et  $v_{max}$ , avec  $v_{max} = 4\sqrt{RT_w} = 953$  m.s<sup>-1</sup> et  $v_{min} = -v_{max}$ .

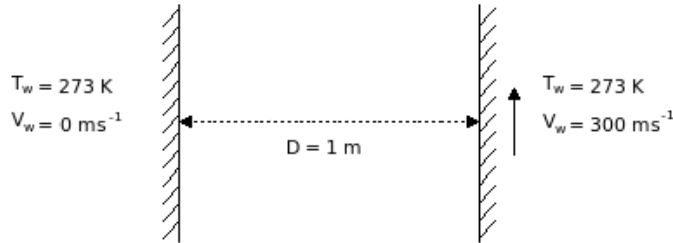


FIGURE 4.8: Configuration de l'écoulement

Le temps de relaxation  $\tau$  s'exprime comme le rapport de la viscosité dynamique et de la pression :

$$\tau = \frac{\mu}{p} \quad \text{avec} \quad \begin{cases} \mu = \mu_0 \left( \frac{T}{T_0} \right)^\omega, \\ p = \rho RT, \end{cases}$$

où  $R$  est la constante spécifique du gaz,  $\mu_0$  et  $T_0$  la viscosité et la température de référence.

Le nombre de Knudsen est défini par :  $\text{Kn} = \frac{\lambda}{D}$ , où le libre parcours moyen des molécules  $\lambda$  est donné par le modèle VHS (voir Bird [14], qui donne des tables pour les valeurs de  $\mu_0$ ,  $T_0$  et  $\omega$ ) :

$$\lambda = \left[ \sqrt{2} \pi d^2 \frac{\rho}{m} \left( \frac{T_0}{T} \right)^{\omega - \frac{1}{2}} \right]^{-1} \quad \text{avec} \quad d = \left[ \frac{5(\alpha + 1)(\alpha + 2) \sqrt{\pi^{-1} m k_b T}}{4\alpha(5 - 2\omega)(7 - 2\omega)\mu_0} \right]^{\frac{1}{2}}$$

où  $k_b$  est la constante de Boltzmann et  $m$  la masse moléculaire du gaz. Les paramètres de l'écoulement et les différentes propriétés du gaz sont données dans le tableau 4.1.

Nature du gaz		Argon	
Température des plaques	273 K	$m$	$0,663.10^{-25}$ kg
Vitesse plaque gauche	$0$ m.s <sup>-1</sup>	$R$	$208,24$ J.kg <sup>-1</sup> .K <sup>-1</sup>
Vitesse plaque droite	$300$ m.s <sup>-1</sup>	$\mu_0$	$2,117.10^{-5}$ Pa.s
Distance entre les plaques	$1$ m	$T_0$	$273,15$ K
		$\omega$	$0,81$
		$\alpha$	$1$
		Kn	$9,25.10^{-3}$

TABLE 4.1: Ecoulement de Couette : paramètres de l'écoulement et propriétés du gaz

On montre sur la figure 4.9 la vitesse horizontale, le flux de chaleur et la pression pour les 4 schémas proposés ici. On n'observe pas de différence significative sur les autres quantités.

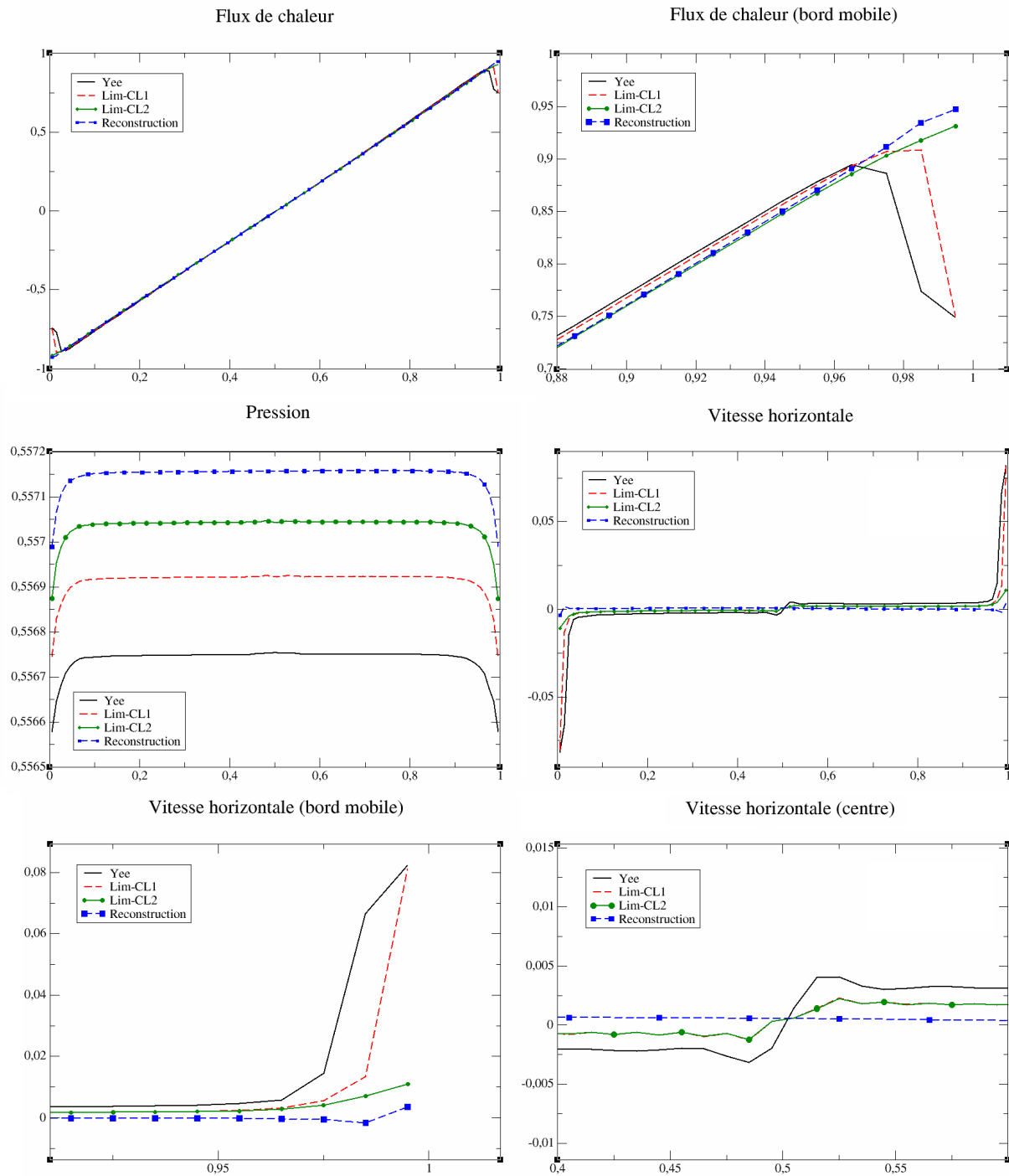


FIGURE 4.9: Vitesse, flux de chaleur et pression obtenus avec les schémas d'ordre 2

On constate que le schéma de Yee et le schéma limité avec conditions au bord d'ordre 1 ont

un très mauvais comportement près des bords où on observe une variation brutale de la vitesse et du flux de chaleur. Ce résultat est cohérent avec l'analyse exposée au paragraphe 4.3.2.2 : le schéma de Yee dégénère à l'ordre 1 à la paroi.

On observe également sur les résultats obtenus avec le schéma de Yee une irrégularité vers le centre du domaine, très visible sur la vitesse, que semble aussi produire le schéma avec limiteur (quelle que soit la condition aux limites). Il n'y a pas d'explication évidente de ce phénomène, qui est sans doute un effet des limiteurs, sans lien avec les conditions aux limites.

Enfin, on peut voir que le schéma de Yee donne une pression un peu plus basse que les schémas à reconstruction linéaire.

Le schéma à reconstruction non limitée semble le meilleur des trois. Il est donc important d'avoir des conditions aux limites d'ordre 2, ce qui n'est possible qu'en utilisant un schéma à reconstruction linéaire, et non un limiteur de flux ; la limitation de pente semble avoir un effet plus délétère qu'autre chose, sans que la raison en soit claire. **Dans toute la suite, on utilisera donc le schéma à reconstruction non limitée à l'ordre 2.**

### 4.3.3 Schéma d'ordre 4

L'extension des méthodes Volumes Finis aux ordres supérieurs à 2 est généralement complexe. On présente ici un schéma d'ordre 4 introduit par Yamamoto et Daiguji dans [72], puis repris par Berthon et Marche dans [12] pour les équations de Saint-Venant et par Duclos dans le cadre des équations cinétiques dans [34] afin de bénéficier de certaines propriétés du schéma, notamment la positivité de la solution numérique.

Ce schéma repose sur une reconstruction de la distribution  $f$  aux interfaces du maillage, comme pour les méthodes d'ordre 2 présentées au paragraphe 4.3.2.1, mais cette fois du 4<sup>e</sup> ordre. On écrit de la même façon que précédemment :

$$\mathcal{F}_{i+\frac{1}{2},k}^n = v_k^+ f_{i+\frac{1}{2},k}^{n,-} + v_k^- f_{i+\frac{1}{2},k}^{n,+}.$$

Les valeurs reconstruites aux interfaces s'écrivent désormais

$$\begin{cases} f_{i+\frac{1}{2},k}^{n,-} = f_{i,k}^n + \Delta f_{i,k}^+, \\ f_{i+\frac{1}{2},k}^{n,+} = f_{i+1,k}^n + \Delta f_{i+1,k}^-, \end{cases}$$

avec les incréments de reconstruction

$$\begin{cases} \Delta f_{i,k}^- = -\frac{1}{6} \left( 2\Delta^* \bar{f}_{i-\frac{1}{2},k} + \Delta^* \tilde{f}_{i+\frac{1}{2},k} \right), \\ \Delta f_{i,k}^+ = \frac{1}{6} \left( \Delta^* \bar{f}_{i-\frac{1}{2},k} + 2\Delta^* \tilde{f}_{i+\frac{1}{2},k} \right), \end{cases}$$

où

$$\begin{cases} \Delta^* \bar{f}_{i-\frac{1}{2},k} = \text{minmod} \left( \Delta^* f_{i-\frac{1}{2},k}, b\Delta^* f_{i+\frac{1}{2},k} \right), \\ \Delta^* \tilde{f}_{i+\frac{1}{2},k} = \text{minmod} \left( \Delta^* f_{i+\frac{1}{2},k}, b\Delta^* f_{i-\frac{1}{2},k} \right), \end{cases} \quad (4.30)$$

et

$$\begin{cases} \Delta^* f_{i+\frac{1}{2},k} = \Delta f_{i+\frac{1}{2},k} - \frac{1}{6} \Delta^3 \bar{f}_{i+\frac{1}{2},k}, \\ \Delta^3 f_{i+\frac{1}{2},k} = \Delta \bar{f}_{i-\frac{1}{2},k} - 2\Delta \bar{f}_{i+\frac{1}{2},k} + \Delta \bar{f}_{i+\frac{3}{2},k}, \end{cases}$$

et enfin

$$\begin{cases} \Delta \bar{f}_{i-\frac{1}{2},k} = \minmod \left( \Delta f_{i-\frac{1}{2},k}, 2\Delta f_{i+\frac{1}{2},k}, 2\Delta f_{i+\frac{3}{2},k} \right), \\ \Delta \bar{f}_{i+\frac{1}{2},k} = \minmod \left( \Delta f_{i+\frac{1}{2},k}, 2\Delta f_{i+\frac{3}{2},k}, 2\Delta f_{i-\frac{3}{2},k} \right), \\ \Delta \bar{f}_{i+\frac{3}{2},k} = \minmod \left( \Delta f_{i+\frac{3}{2},k}, 2\Delta f_{i-\frac{1}{2},k}, 2\Delta f_{i+\frac{1}{2},k} \right), \end{cases}$$

avec l'incrément :

$$\Delta f_{i+\frac{1}{2},k} = f_{i+1,k}^n - f_{i,k}^n.$$

Le paramètre  $b$  des relations (4.30) est choisi tel que :

$$1 < b \leq 4.$$

On introduit une limitation afin de respecter la positivité :

$$\begin{cases} \Delta f_{i,k}^{-,lim} = \theta \max \left( \Delta f_{i,k}^-, -f_{i,k}^n \right), \\ \Delta f_{i,k}^{+,lim} = \theta \max \left( \Delta f_{i,k}^+, -f_{i,k}^n \right), \end{cases}$$

avec :

$$\theta = \begin{cases} 1 & \text{si } \sigma \leq 0, \\ \min \left( 1, \sigma^{-1} f_{i,k}^n \right) & \text{sinon,} \end{cases} \quad \sigma = \max \left( \Delta f_i^-, -f_{i,k}^n \right) + \max \left( \Delta f_i^+, -f_{i,k}^n \right).$$

Il n'y a pas a priori de façon simple de construire des conditions aux limites d'ordre élevé sur ce schéma fortement non linéaire. On se contente donc de conditions d'ordre 1, comme pour les schémas d'ordre 2 à limiteurs de flux décrits au paragraphe 4.3.2.2.

## 4.4 Schéma Galerkin Discontinu

On cherche maintenant à appliquer une discrétisation de type Galerkin Discontinu avec des éléments linéaires sur chaque maille au modèle (4.8).

On va considérer uniquement l'équation sur la distribution  $f$ , celles sur  $g$  et  $h$  se traitant naturellement de la même façon :

$$\frac{\partial f_k}{\partial t} + v_k \frac{\partial f_k}{\partial x} = \frac{1}{\tau} (\mathcal{M}_k[\mathbf{U}] - f_k), \quad \forall k \in \mathcal{K}. \quad (4.31)$$

Au lieu d'utiliser la formulation variationnelle de la méthode, présentée au chapitre 3, adaptée à son analyse mathématique, nous préférons utiliser ici une formulation équivalente, plus proche de la méthode Volumes Finis, dont l'implémentation est immédiate.

#### 4.4.1 Ecriture du schéma semi-discret

Dans un premier temps, on ne considèrera pas la discrétisation en temps, qui sera abordée au paragraphe 4.4.2.

On choisit pour fonctions de base sur chaque maille les fonctions linéaires :

$$\begin{cases} \varphi_{i,1}(\xi) = 1 - \xi, \\ \varphi_{i,2}(\xi) = \xi \end{cases},$$

où  $\xi$  est l'abscisse locale sur chaque maille  $C_i$  :  $\xi = \frac{x - x_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}} \in [0; 1]$ , avec  $C_i = [x_{i-\frac{1}{2}}; x_{i+\frac{1}{2}}]$ .

On projette dans chaque maille  $C_i$  la fonction  $f_k$  sur la base  $(\varphi_{i,1}, \varphi_{i,2})$  de la façon suivante :

$$f_{k|C_i}(x) = f_{i,k}(x) = \sum_{p=1}^2 f_{i,k,p} \cdot \varphi_{i,p}(x). \quad (4.32)$$

Avec les fonction de base choisies,  $f_{i,k,1}$  et  $f_{i,k,2}$  sont respectivement les limites de  $f_k$  à gauche et à droite de la maille  $C_i$ .

Dans l'équation instationnaire, les coefficients  $f_{i,k,p}$  dépendent de  $t$ , ce qui revient à faire une séparation des variables de  $f_k$ . On n'écrira pas ici la dépendance en  $t$  de ce coefficient car on cherche une solution stationnaire, on verra cependant apparaître une dérivée temporelle "factice" pour la marche en temps.

On réinjecte (4.32) dans l'équation (4.31) et on projette sur une fonction de base  $\varphi_{i,q}$  ( $q = 1$  ou  $2$ ) :

$$\int_{C_i} \sum_p \left( \frac{\partial f_{i,k,p}}{\partial t} \varphi_{i,p}(x) + v_k f_{i,k,p} \frac{\partial \varphi_{i,p}(x)}{\partial x} \right) \varphi_{i,q}(x) dx = \int_{C_i} \frac{1}{\tau_i} (\mathcal{M}_k[\mathbf{U}_i] - f_{i,k}) \varphi_{i,q}(x) dx, \quad (4.33)$$

soit :

$$\sum_p \int_{C_i} \frac{\partial f_{i,k,p}}{\partial t} \varphi_{i,p}(x) \varphi_{i,q}(x) dx + \sum_p \int_{C_i} v_k f_{i,k,p} \frac{\partial \varphi_{i,p}(x)}{\partial x} \varphi_{i,q}(x) dx = \int_{C_i} \frac{1}{\tau_i} (\mathcal{M}_k[\mathbf{U}_i] - f_{i,k}) \varphi_{i,q}(x) dx.$$

On développe maintenant précisément chaque terme de cette équation.

##### 4.4.1.1 Terme instationnaire

Ce terme se calcule simplement :

$$\sum_p \int_{C_i} \frac{\partial f_{i,k,p}}{\partial t} \varphi_{i,p}(x) \varphi_{i,q}(x) dx = \sum_p m_{qp}^i \frac{\partial f_{i,k,p}}{\partial t}, \text{ avec } m_{qp}^i = \int_{C_i} \varphi_{i,p}(x) \varphi_{i,q}(x) dx.$$

On définit la matrice " de masse" locale :

$$\bar{M}_i = (m_{qp}^i)_{1 \leq q,p \leq 2} = \frac{\Delta x_i}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \quad (4.34)$$

#### 4.4.1.2 Terme de flux

Le second terme est intégré par parties :

$$\begin{aligned}
\sum_p \left[ \int_{C_i} v_k \frac{\partial \varphi_{i,p}(x)}{\partial x} \varphi_{i,q}(x) dx \right] f_{i,k,p} \\
&= \sum_p \left[ -v_k \int_{C_i} \varphi_{i,p}(x) \frac{\partial \varphi_{i,q}(x)}{\partial x} dx + \left[ v_k \varphi_{i,p}(x) \varphi_{i,q}(x) \right]_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \right] f_{i,k,p} \\
&= -v_k \int_{C_i} \sum_p f_{i,k,p} \cdot \varphi_{i,p}(x) \frac{\partial \varphi_{i,q}(x)}{\partial x} dx + \left[ \sum_p v_k f_{i,k,p} \cdot \varphi_{i,p}(x) \varphi_{i,q}(x) \right]_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \\
&= -v_k \int_{C_i} f_{i,k}(x) \frac{\partial \varphi_{i,q}(x)}{\partial x} dx + \left[ v_k f_{i,k}(x) \varphi_{i,q}(x) \right]_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}.
\end{aligned}$$

Le terme intégral s'écrit :

$$\begin{aligned}
-v_k \int_{C_i} f_{i,k}(x) \frac{\partial \varphi_{i,q}(x)}{\partial x} dx &= -v_k \int_{C_i} f_{i,k}(x) \frac{(-1)^q}{\Delta x_i} dx = -v_k \sum_p \frac{(-1)^q}{\Delta x} f_{i,k,p} \int_{C_i} \varphi_{i,p}(x) dx \\
&= -v_k \sum_p \frac{(-1)^q}{\Delta x_i} f_{i,k,p} \frac{\Delta x_i}{2} \\
&= \sum_p \left( \frac{(-1)^{q+1}}{2} v_k \right) f_{i,k,p} = \sum_p d_{qp} f_{i,k,p}.
\end{aligned}$$

On définit alors la matrice :

$$\bar{\bar{D}}_k = (d_{qp})_{1 \leq q, p \leq 2} = \frac{v_k}{2} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}. \quad (4.35)$$

Pour échanger l'information entre les mailles, il est important de définir le terme de bord par un flux décentré, comme pour les Volumes Finis :

$$\left[ v_k f_{i,k}(x) \varphi_{i,q}(x) \right]_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} = \mathcal{F}_{i+\frac{1}{2},q} - \mathcal{F}_{i-\frac{1}{2},q},$$

avec

$$\begin{aligned}
\mathcal{F}_{i+\frac{1}{2},q} &= v_k^+ \cdot f_{i,k}(x_{i+\frac{1}{2}}) \varphi_{i,q}(x_{i+\frac{1}{2}}) + v_k^- \cdot f_{i+1,k}(x_{i+\frac{1}{2}}) \varphi_{i,q}(x_{i+\frac{1}{2}}) \\
&= v_k^+ \cdot f_{i,k,2} \varphi_{i,q}(x_{i+\frac{1}{2}}) + v_k^- \cdot f_{i+1,k,1} \varphi_{i,q}(x_{i+\frac{1}{2}}), \\
\mathcal{F}_{i-\frac{1}{2},q} &= v_k^+ \cdot f_{i-1,k}(x_{i-\frac{1}{2}}) \varphi_{i,q}(x_{i-\frac{1}{2}}) + v_k^- \cdot f_{i,k}(x_{i-\frac{1}{2}}) \varphi_{i,q}(x_{i-\frac{1}{2}}) \\
&= v_k^+ \cdot f_{i-1,k,2} \varphi_{i,q}(x_{i-\frac{1}{2}}) + v_k^- \cdot f_{i,k,1} \varphi_{i,q}(x_{i-\frac{1}{2}}).
\end{aligned}$$

Ce qui donne :

$$\overrightarrow{\mathcal{F}_{i+\frac{1}{2}}} = \begin{pmatrix} \mathcal{F}_{i+\frac{1}{2},1} \\ \mathcal{F}_{i+\frac{1}{2},2} \end{pmatrix} = \begin{pmatrix} 0 \\ v_k^+ \cdot f_{i,k,2} + v_k^- \cdot f_{i+1,k,1} \end{pmatrix}, \quad \overrightarrow{\mathcal{F}_{i-\frac{1}{2}}} = \begin{pmatrix} \mathcal{F}_{i-\frac{1}{2},1} \\ \mathcal{F}_{i-\frac{1}{2},2} \end{pmatrix} = \begin{pmatrix} v_k^+ \cdot f_{i-1,k,2} + v_k^- \cdot f_{i,k,1} \\ 0 \end{pmatrix}.$$

#### 4.4.1.3 Terme de collision

On doit calculer le terme  $\int_{C_i} \frac{1}{\tau_i} (\mathcal{M}_k[\mathbf{U}_i] - f_{i,k}) \varphi_{i,q}(x) dx$ . La façon naturelle de procéder serait d'exprimer la dépendance de  $\frac{1}{\tau_i}$  et  $\mathcal{M}_k[\mathbf{U}_i]$  par rapport à  $f$  et de projeter ensuite  $f$  sur la base locale, mais cette dépendance est beaucoup trop complexe. On projette donc directement les quantités  $\frac{1}{\tau_i} \mathcal{M}_k[\mathbf{U}_i]$  et  $\frac{1}{\tau_i} f_{i,k}$  sur la base :

$$\frac{1}{\tau_i} \mathcal{M}_k[\mathbf{U}_i] = \sum_p \left( \frac{\mathcal{M}}{\tau} \right)_{i,k,p} \varphi_{i,p}(x), \quad \frac{1}{\tau_i} f_{i,k} = \sum_p \left( \frac{f}{\tau} \right)_{i,k,p} \varphi_{i,p}(x).$$

Les composantes  $\left( \frac{\mathcal{M}}{\tau} \right)_{i,k,p}$  et  $\left( \frac{f}{\tau} \right)_{i,k,p}$  étant, compte-tenu de la base choisie, les valeurs ponctuelles des grandeurs  $\frac{1}{\tau_i} \mathcal{M}_k[\mathbf{U}_i]$  et  $\frac{1}{\tau_i} f_{i,k}$  à gauche et à droite de la cellule  $C_i$ , on a clairement

$$\left( \frac{\mathcal{M}}{\tau} \right)_{i,k,p} = \frac{\mathcal{M}_{i,k,p}}{\tau_{i,p}}, \quad \left( \frac{f}{\tau} \right)_{i,k,p} = \frac{f_{i,k,p}}{\tau_{i,p}}$$

en appliquant la même approximation sur  $\tau_i$  et  $\mathcal{M}_k[\mathbf{U}_i]$  :

$$\tau_i = \sum_p \tau_{i,p} \varphi_{i,p}(x), \quad \mathcal{M}_k[\mathbf{U}_i] = \sum_p \mathcal{M}_{i,k,p} \varphi_{i,p}(x).$$

Ceci donne finalement :

$$\begin{aligned} \int_{C_i} \frac{1}{\tau_i} (\mathcal{M}_k[\mathbf{U}_i] - f_{i,k}) \varphi_{i,q}(x) dx &= \int_{C_i} \left[ \sum_p \frac{1}{\tau_{i,p}} (\mathcal{M}_{i,k,p} - f_{i,k,p}) \varphi_{i,p}(x) \right] \varphi_{i,q}(x) dx \\ &= \sum_p \frac{1}{\tau_{i,p}} (\mathcal{M}_{i,k,p} - f_{i,k,p}) \int_{C_i} \varphi_{i,p}(x) \varphi_{i,q}(x) dx \\ &= \sum_p m_{qp}^i \frac{1}{\tau_{i,p}} (\mathcal{M}_{i,k,p} - f_{i,k,p}). \end{aligned}$$

Il reste à exprimer les composantes  $\mathcal{M}_{i,k,p}$ . On calcule la distribution d'équilibre discrète en utilisant les variables entropiques :

$$\mathcal{M}_k[\mathbf{U}_i] = \exp(\boldsymbol{\alpha}_i \cdot \mathbf{m}_k),$$

où  $\mathbf{m}_k = (1, v_k, \frac{1}{2}|v_k|^2)^T$ , et  $\boldsymbol{\alpha}_i$  est la solution du problème

$$\sum_k \mathbf{m}_k \cdot \exp(\mathbf{m}_k \cdot \boldsymbol{\alpha}_i(x)) \Delta v_k = \boldsymbol{\rho}_i(x), \quad (4.36)$$

avec  $\boldsymbol{\rho}_i$  le vecteur des moments discrets de  $f$  :

$$\boldsymbol{\rho}_i(x) = \sum_k \mathbf{m}_k f_{i,k}(x) \Delta v_k = \sum_k \mathbf{m}_k \sum_p f_{i,k,p} \varphi_{i,p}(x) \Delta v_k = \sum_p \underbrace{\sum_k \mathbf{m}_k f_{i,k,p} \Delta v_k}_{\boldsymbol{\rho}_{i,p}} \varphi_{i,p}(x).$$



On obtient ainsi naturellement :

$$\begin{cases} \rho_{i,1} = \sum_k \mathbf{m}_k f_{i,k}(x_{i-\frac{1}{2}}) \Delta v_k, \\ \rho_{i,2} = \sum_k \mathbf{m}_k f_{i,k}(x_{i+\frac{1}{2}}) \Delta v_k. \end{cases}$$

De même, on a :

$$\begin{cases} \mathcal{M}_{i,k,1} = \mathcal{M}_{i,k}(x_{i-\frac{1}{2}}) = \exp(\mathbf{m}_k \cdot \boldsymbol{\alpha}_i(x_{i-\frac{1}{2}})), \\ \mathcal{M}_{i,k,2} = \mathcal{M}_{i,k}(x_{i+\frac{1}{2}}) = \exp(\mathbf{m}_k \cdot \boldsymbol{\alpha}_i(x_{i+\frac{1}{2}})). \end{cases}$$

Les composantes de l'équilibre discret  $\mathcal{M}_{i,k}$  sur la base  $(\varphi_{i,1}, \varphi_{i,2})$  s'obtiennent donc simplement en résolvant le problème (4.36) en  $x_{i-\frac{1}{2}}$  et  $x_{i+\frac{1}{2}}$ .

#### 4.4.1.4 Formulation matricielle

En posant :

$$\mathbf{f}_{i,k} = (f_{i,k,p})_p, \quad \boldsymbol{\varepsilon}_{i,k} = \left( \frac{1}{\tau_{i,p}} \mathcal{M}_{i,k,p} \right)_p = (\varepsilon_{i,k,p})_p, \quad \tilde{\mathbf{f}}_{i,k} = \left( \frac{1}{\tau_{i,p}} f_{i,k,p} \right)_p = (\tilde{f}_{i,k,p})_p,$$

l'équation (4.33) écrite pour  $\varphi_{i,q} = \varphi_{i,1}$  puis  $\varphi_{i,2}$  donne le système

$$\bar{M}_i \frac{\partial \mathbf{f}_{i,k}}{\partial t} + \bar{D}_k \mathbf{f}_{i,k} + \mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2} = \bar{M}_i (\boldsymbol{\varepsilon}_{i,k} - \tilde{\mathbf{f}}_{i,k}). \quad (4.37)$$

Les vecteurs flux s'écrivent :

$$\mathcal{F}_{i+1/2} = \begin{pmatrix} \mathcal{F}_{i+\frac{1}{2},1} \\ \mathcal{F}_{i+\frac{1}{2},2} \end{pmatrix} = \begin{pmatrix} 0 \\ v_k^+ \cdot f_{i,k,2} + v_k^- \cdot f_{i+1,k,1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & v_k^+ \end{pmatrix} \mathbf{f}_{i,k} + \begin{pmatrix} 0 & 0 \\ v_k^- & 0 \end{pmatrix} \mathbf{f}_{i+1,k},$$

$$\mathcal{F}_{i-1/2} = \begin{pmatrix} \mathcal{F}_{i-\frac{1}{2},1} \\ \mathcal{F}_{i-\frac{1}{2},2} \end{pmatrix} = \begin{pmatrix} v_k^+ \cdot f_{i-1,k,2} + v_k^- \cdot f_{i,k,1} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & v_k^+ \\ 0 & 0 \end{pmatrix} \mathbf{f}_{i-1,k} + \begin{pmatrix} v_k^- & 0 \\ 0 & 0 \end{pmatrix} \mathbf{f}_{i,k},$$

ce qui donne :

$$\begin{aligned} \mathcal{F}_{i+1/2} - \mathcal{F}_{i-1/2} &= \begin{pmatrix} 0 & -v_k^+ \\ 0 & 0 \end{pmatrix} \mathbf{f}_{i-1,k} + \begin{pmatrix} -v_k^- & 0 \\ 0 & v_k^+ \end{pmatrix} \mathbf{f}_{i,k} + \begin{pmatrix} 0 & 0 \\ v_k^- & 0 \end{pmatrix} \mathbf{f}_{i+1,k} \\ &= \bar{A} \mathbf{f}_{i-1,k} + \bar{B} \mathbf{f}_{i,k} + \bar{C} \mathbf{f}_{i+1,k}. \end{aligned}$$

On multiplie (4.37) par l'inverse de la "matrice de masse"  $\bar{M}_i^{-1} = \frac{2}{\Delta x_i} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$  :

$$\begin{cases} \bar{M}_i^{-1} \bar{D} = \frac{1}{\Delta x_i} \begin{pmatrix} 3v_k & 3v_k \\ -3v_k & -3v_k \end{pmatrix}, & \begin{cases} \bar{M}_i^{-1} \bar{B} = \frac{2}{\Delta x_i} \begin{pmatrix} -2v_k^- & -v_k^+ \\ v_k^- & 2v_k^+ \end{pmatrix}, \\ \bar{M}_i^{-1} \bar{C} = \frac{2}{\Delta x_i} \begin{pmatrix} -v_k^- & 0 \\ 2v_k^- & 0 \end{pmatrix}, \end{cases} \\ \bar{M}_i^{-1} \bar{A} = \frac{2}{\Delta x_i} \begin{pmatrix} 0 & -2v_k^+ \\ 0 & v_k^+ \end{pmatrix}, \end{cases}$$

On obtient finalement l'équation semi-discrète :

$$\begin{aligned} \frac{\partial \mathbf{f}_{i,k}}{\partial t} = & -\frac{1}{\Delta x_i} \begin{pmatrix} 3v_k & 3v_k \\ -3v_k & -3v_k \end{pmatrix} \mathbf{f}_{i,k} \\ & -\frac{2}{\Delta x_i} \left[ \begin{pmatrix} 0 & -2v_k^+ \\ 0 & v_k^+ \end{pmatrix} \mathbf{f}_{i-1,k} + \begin{pmatrix} -2v_k^- & -v_k^+ \\ v_k^- & 2v_k^+ \end{pmatrix} \mathbf{f}_{i,k} + \begin{pmatrix} -v_k^- & 0 \\ 2v_k^- & 0 \end{pmatrix} \mathbf{f}_{i+1,k} \right] \\ & + (\mathcal{E}_{i,k} - \tilde{\mathbf{f}}_{i,k}), \end{aligned} \quad (4.38)$$

soit pour un maillage de pas constant ( $\forall i, \Delta x_i = \Delta x$ ) :

$$\begin{aligned} \frac{\partial}{\partial t} \begin{pmatrix} f_{1,k,1} \\ f_{1,k,2} \\ f_{2,k,1} \\ f_{2,k,2} \\ \vdots \\ f_{(N-1),k,1} \\ f_{(N-1),k,2} \\ f_{N,k,1} \\ f_{N,k,2} \end{pmatrix} = & -\frac{3}{\Delta x} \begin{pmatrix} v_k & v_k & & & & & & & \\ -v_k & -v_k & & & & & & & \\ & & v_k & v_k & & & & & \\ & & -v_k & -v_k & & & & & \\ & & & \ddots & \ddots & \ddots & & & \\ & & & & v_k & v_k & & & \\ & & & & -v_k & -v_k & & & \\ & & & & & & v_k & v_k & \\ & & & & & & -v_k & -v_k & \\ & & & & & & & v_k & v_k \\ & & & & & & & -v_k & -v_k \end{pmatrix} \begin{pmatrix} f_{1,k,1} \\ f_{1,k,2} \\ f_{2,k,1} \\ f_{2,k,2} \\ \vdots \\ f_{(N-1),k,1} \\ f_{(N-1),k,2} \\ f_{N,k,1} \\ f_{N,k,2} \end{pmatrix} \\ & -\frac{2}{\Delta x} \begin{pmatrix} -2v_k^- & -v_k^+ & -v_k^- & & & & & & \\ v_k^- & 2v_k^+ & 2v_k^- & & & & & & \\ & -2v_k^+ & -2v_k^- & -v_k^+ & -v_k^- & & & & \\ & v_k^+ & v_k^- & 2v_k^+ & 2v_k^- & & & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & & -2v_k^+ & 2v_k^- & -v_k^+ & -v_k^- & & \\ & & & v_k^+ & v_k^- & 2v_k^+ & 2v_k^- & -v_k^+ & -v_k^- \\ & & & & & -2v_k^+ & -2v_k^- & -v_k^+ & -v_k^- \\ & & & & & v_k^+ & v_k^- & 2v_k^+ & 2v_k^- \end{pmatrix} \begin{pmatrix} f_{1,k,1} \\ f_{1,k,2} \\ f_{2,k,1} \\ f_{2,k,2} \\ \vdots \\ f_{(N-1),k,1} \\ f_{(N-1),k,2} \\ f_{N,k,1} \\ f_{N,k,2} \end{pmatrix} \\ & + \begin{pmatrix} \mathcal{E}_{1,k,1} \\ \mathcal{E}_{1,k,2} \\ \mathcal{E}_{2,k,1} \\ \mathcal{E}_{2,k,2} \\ \vdots \\ \mathcal{E}_{(N-1),k,1} \\ \mathcal{E}_{(N-1),k,2} \\ \mathcal{E}_{N,k,1} \\ \mathcal{E}_{N,k,2} \end{pmatrix} - \begin{pmatrix} \tilde{f}_{1,k,1} \\ \tilde{f}_{1,k,2} \\ \tilde{f}_{2,k,1} \\ \tilde{f}_{2,k,2} \\ \vdots \\ \tilde{f}_{(N-1),k,1} \\ \tilde{f}_{(N-1),k,2} \\ \tilde{f}_{N,k,1} \\ \tilde{f}_{N,k,2} \end{pmatrix}. \end{aligned}$$

#### 4.4.2 Discrétisation temporelle

Les schémas d'Euler explicites sont instables avec Galerkin Discontinu (voir [27]). On utilise donc pour le schéma explicite la discrétisation Runge-Kutta 2 "point milieu" (Heun) : en notant

$$\frac{\partial f_k}{\partial t} = \mathcal{R}(t, f_k),$$

on a :

$$f_k^{n+1} = f_k^n + \Delta t \mathcal{R}\left(t + \frac{\Delta t}{2}, f_k^n + \frac{\Delta t}{2} \mathcal{R}(t, f_k^n)\right).$$

Le schéma implicite est celui décrit au paragraphe 4.2.

#### 4.4.3 Conditions aux limites

Le schéma GD est compact :  $f_{i,k}^{n+1}$  est une fonction de  $f_{i,k}^n$ ,  $f_{i-1,k,2}^n$  et  $f_{i+1,k,1}^n$ . On n'a donc besoin que d'une seule cellule fictive à la paroi, où on peut ne définir  $f_k$  que sur le bord correspondant à l'interface avec le gaz, c'est-à-dire  $f_{0,k,2}$  et  $f_{\text{imaxp1},k,1}$ .

On impose une réflexion diffuse du gaz à la paroi. En notant  $u_g$  et  $T_g$  la vitesse et la température de la paroi gauche et  $u_d$  et  $T_d$  celles de la paroi droite, on définit  $f_{0,k,2}$  et  $f_{\text{imaxp1},k,1}$  par :

$$\begin{cases} f_{0,k,2} = \sigma_g \mathcal{M}_k[1, u_g, T_g], & v_k > 0, \\ f_{\text{imaxp1},k,1} = \sigma_d \mathcal{M}_k[1, u_d, T_d], & v_k < 0. \end{cases}$$

Le paramètre  $\sigma_g$  est défini par :

$$\sigma_g = - \left( \sum_k v_k^- f_{1,k,1} \Delta v_k \right) \left( \sum_k v_k^+ \mathcal{M}_k[1, u_g, T_g] \Delta v_k \right)^{-1},$$

et de même :

$$\sigma_d = - \left( \sum_k v_k^- f_{\text{imax},k,2} \Delta v_k \right) \left( \sum_k v_k^+ \mathcal{M}_k[1, u_d, T_d] \Delta v_k \right)^{-1}.$$

Le schéma Galerkin Discontinu a donc des conditions aux limites semblables à celles du schéma Volumes Finis d'ordre 1. Cependant, son ordre est plus élevé : le schéma Galerkin Discontinu avec un flux upwind est théoriquement d'ordre 3/2 (voir les références données dans [27]).

### 4.5 Comparaison des schémas Volumes Finis et Galerkin Discontinu

On compare les schémas volumes finis d'ordre 1, 2 et 4 (notés respectivement VF1, VF2 et VF4) avec le schéma Galerkin Discontinu (noté GD) sur le cas décrit au paragraphe 4.3.2.3 : écoulement de Couette pour une vitesse de plaque mobile de 300 m.s<sup>-1</sup>. Tous les schémas sont implicites, suivant la structure exposée au paragraphe 4.2 ; la CFL maximale est fixée à 10000.

Le résidu  $R^n$  est défini comme la norme  $L^2$  du second membre  $\mathbf{RHS}^n$  :

$$R^n = \left( \sum_i \sum_k \left[ \left( v_k \frac{\partial f_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{M}_k[\mathbf{U}_i^n] - f_{i,k}^n) \right)^2 + \left( v_k \frac{\partial g_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{N}_k[\mathbf{U}_i^n] - g_{i,k}^n) \right)^2 + \left( v_k \frac{\partial h_{i,k}^n}{\partial x} + \frac{1}{\tau_i^n} (\mathcal{P}_k[\mathbf{U}_i^n] - h_{i,k}^n) \right)^2 \right] \Delta x_i \Delta v_k \right)^{1/2}.$$

On teste les schémas d'abord sur des maillages réguliers, afin de déterminer expérimentalement leurs ordres de convergence. On passe ensuite à des maillages raffinés en zone de proche paroi, afin de tester les différents schémas sur des maillages qui sont utilisés en aérodynamique.

#### 4.5.1 Maillages réguliers : estimation de l'erreur et ordres de convergence

On utilise des maillages de 12, 100, 800 et 6400 mailles.

Le critère de convergence (diminution de 5 décades du résidu) est atteint en environ 1250 itérations pour le schéma d'ordre 1, environ 1600 pour les trois autres :

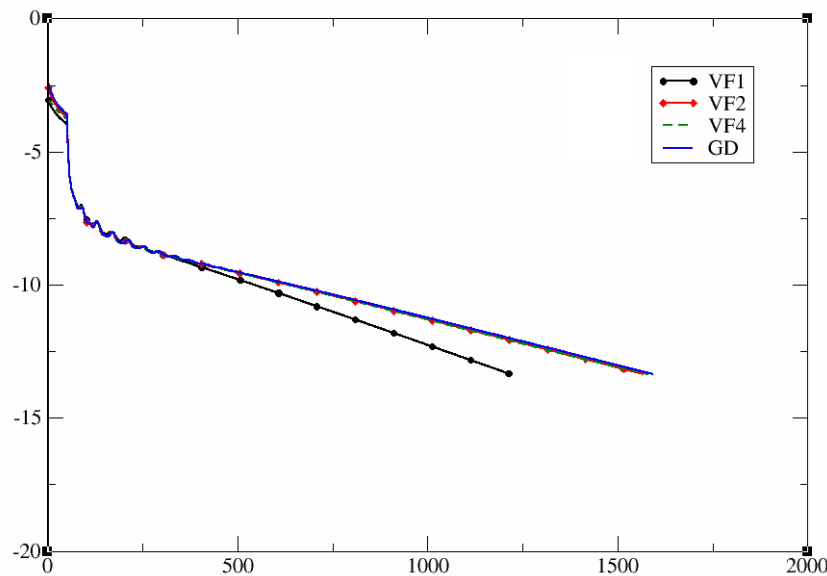
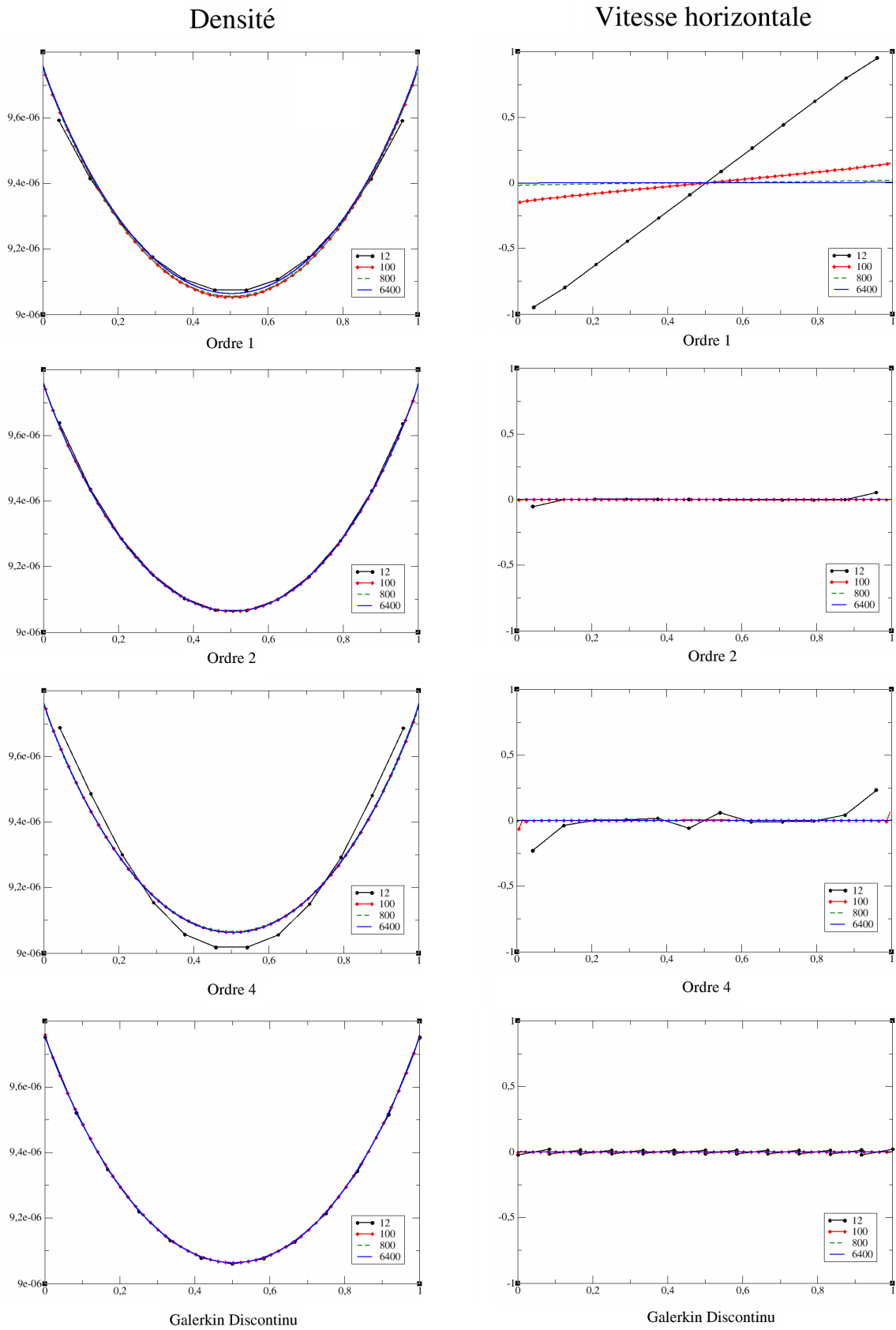


FIGURE 4.10: Evolution du résidu (échelle logarithmique) en fonction du nombre d'itérations

On présente sur les figures 4.11, 4.12 et 4.13 la densité, les vitesses, la température, la pression et le flux de chaleur obtenus pour chaque schéma sur les différents maillages.

On observe que le schéma GD est très proche du schéma VF2, aussi bien pour les résultats donnés que pour son évolution vers la convergence. Les résultats sont beaucoup plus précis que ceux du schéma VF1, ce qui se voit particulièrement sur la vitesse, la température et la pression.

Le schéma d'ordre 4 donne de mauvais résultats, dus à ses conditions aux limites inadéquates.

FIGURE 4.11: *Densité et vitesse horizontale*

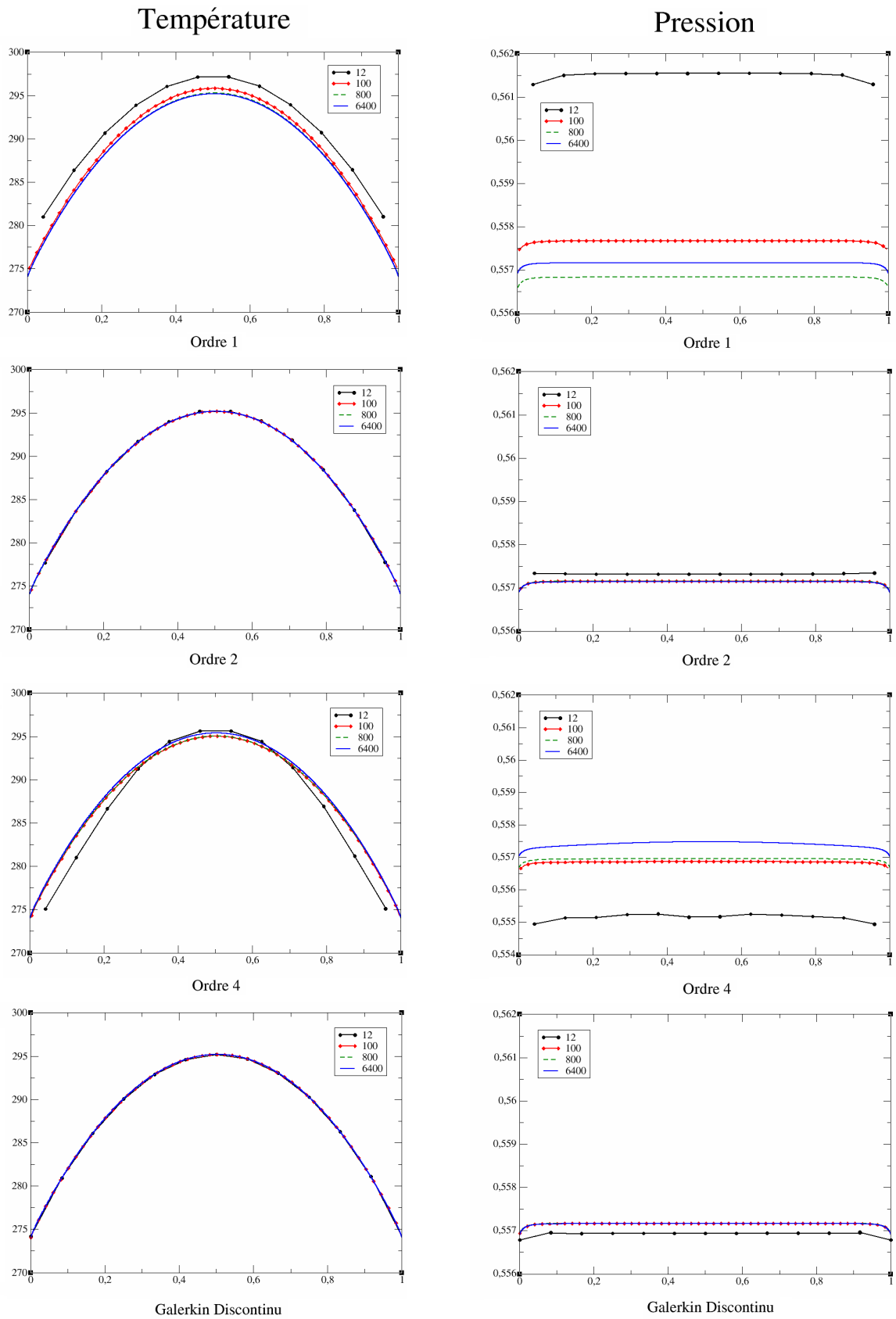


FIGURE 4.12: Température et pression

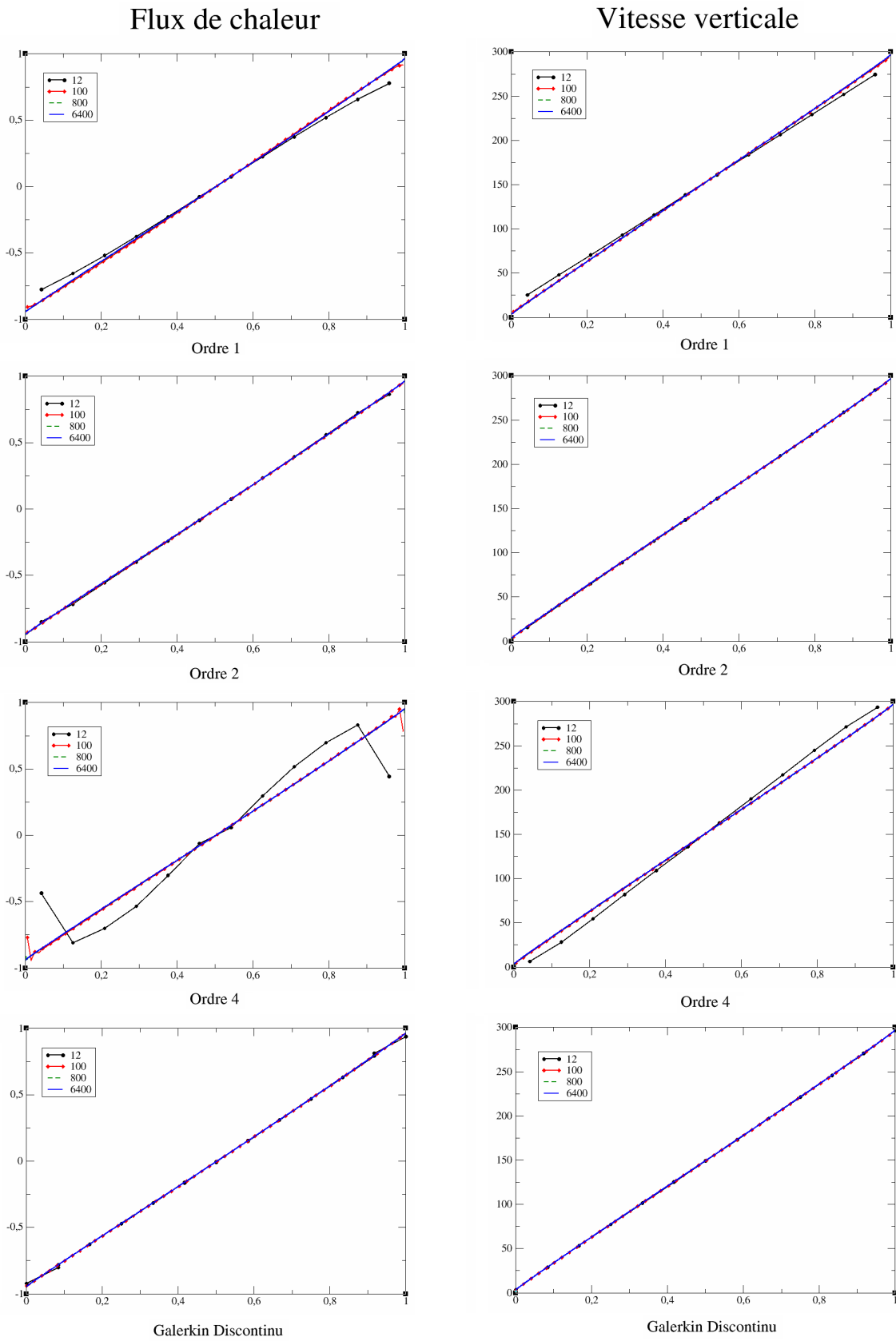


FIGURE 4.13: Flux de chaleur et vitesse verticale

A partir des résultats obtenus, on calcule les ordres de convergence des schémas VF1, VF2 et GD sur différentes grandeurs ; on ne considère pas le schéma VF4, dont on a vu les mauvais résultats. On donne également les courbes obtenues avec le schéma à limiteur de flux (noté Yee et décrit à la section 4.3.2.2).

Les ordres de convergence sont calculés de la façon suivante. La solution analytique du problème étant inconnue, on prend comme solution de référence pour une grandeur donnée  $u$  la solution numérique calculée sur le maillage le plus fin (soit 6400 mailles dans notre cas). On note cette solution  $\bar{u}$ . Si  $u$  est la solution numérique obtenue sur un maillage  $T$  de pas  $\Delta x$ , on définit sur chaque maille  $C_i$  de  $T$  l'écart  $\Delta_i$ , entre les deux solutions :

$$\Delta_i = u_i - \bar{u}(x_i)$$

où  $x_i$  est le centre de la maille  $C_i$ . La fonction  $\bar{u}$  étant discrète, la valeur  $\bar{u}(x_i)$  est obtenue par interpolation si nécessaire.

☞ NB : rigoureusement, on devrait plutôt évaluer

$$\Delta_i = \bar{u}_i - \frac{1}{\Delta x_i} \int_{C_i} \bar{u}(x) dx$$

mais pour des ordres de convergence inférieurs à 2, la première formule est suffisante car

$$\frac{1}{\Delta x_i} \int_{C_i} \bar{u}(x) dx = \bar{u}(x_i) + \mathcal{O}(\Delta x_i^2). \quad \square$$

On définit les erreurs, respectivement en norme  $L^\infty$  et en norme  $L^2$  par

$$E_{\Delta x}^\infty = \max_i |\Delta_i|,$$

et

$$E_{\Delta x}^2 = \left( \sum_i (\Delta_i)^2 \Delta x_i \right)^{1/2}.$$

En supposant un comportement de l'erreur de la forme

$$E_{\Delta x} = C \Delta x^p$$

où  $C > 0$  est une constante, et  $p$  l'ordre de convergence du schéma, on détermine  $p$  en calculant l'erreur sur deux maillages distincts de pas respectifs  $\Delta x_1$  et  $\Delta x_2$ . L'ordre du schéma est alors la pente de la droite représentant l'erreur en fonction du pas de maillage en échelle logarithmique :

$$p = \frac{\log(E_{\Delta x_1}) - \log(E_{\Delta x_2})}{\log(\Delta x_1) - \log(\Delta x_2)}$$

Ces courbes sont représentées pour différentes grandeurs sur les figures 4.14 et 4.15, les ordres correspondants sont donnés dans le tableau 4.2.



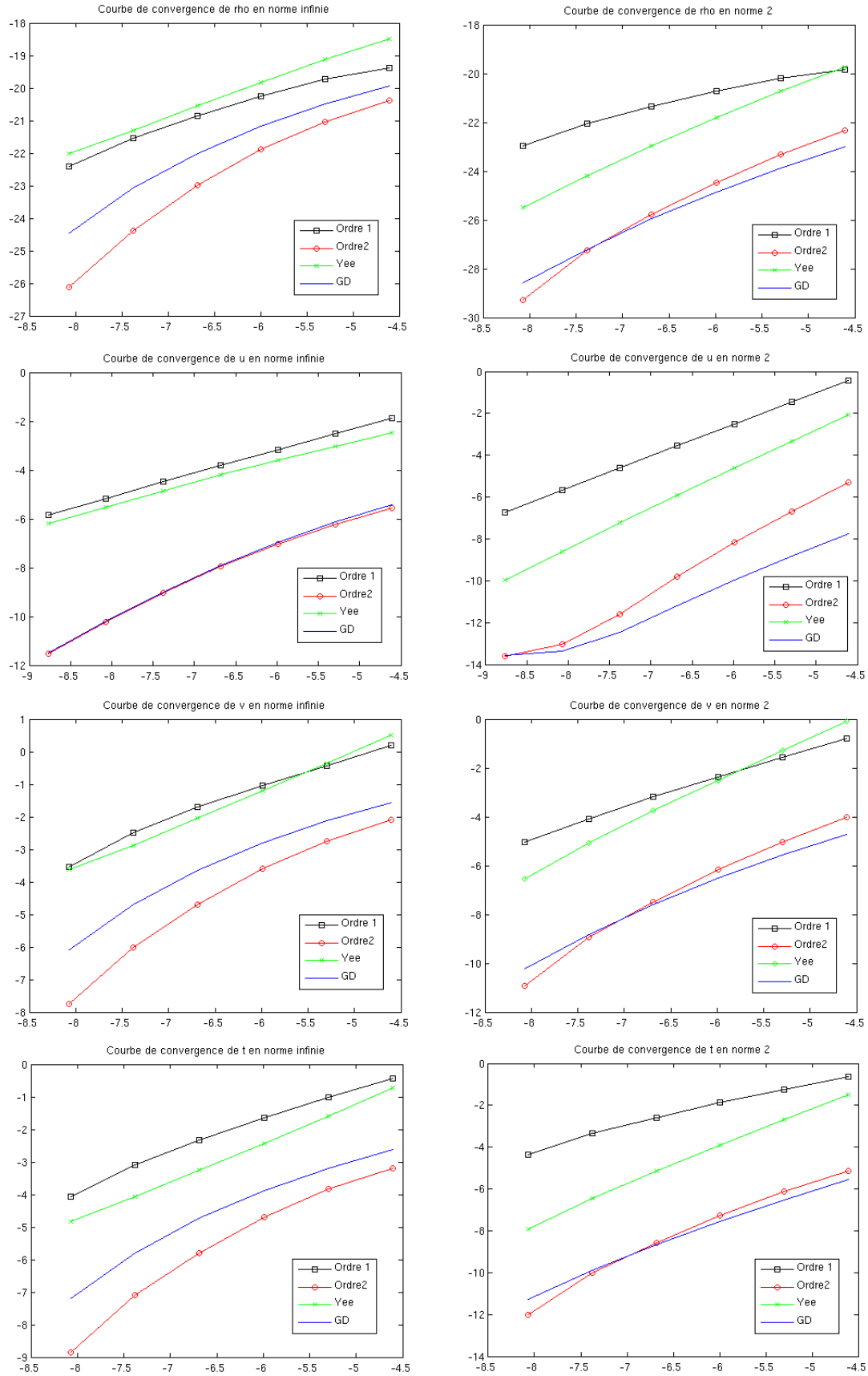


FIGURE 4.14: Courbes de convergence

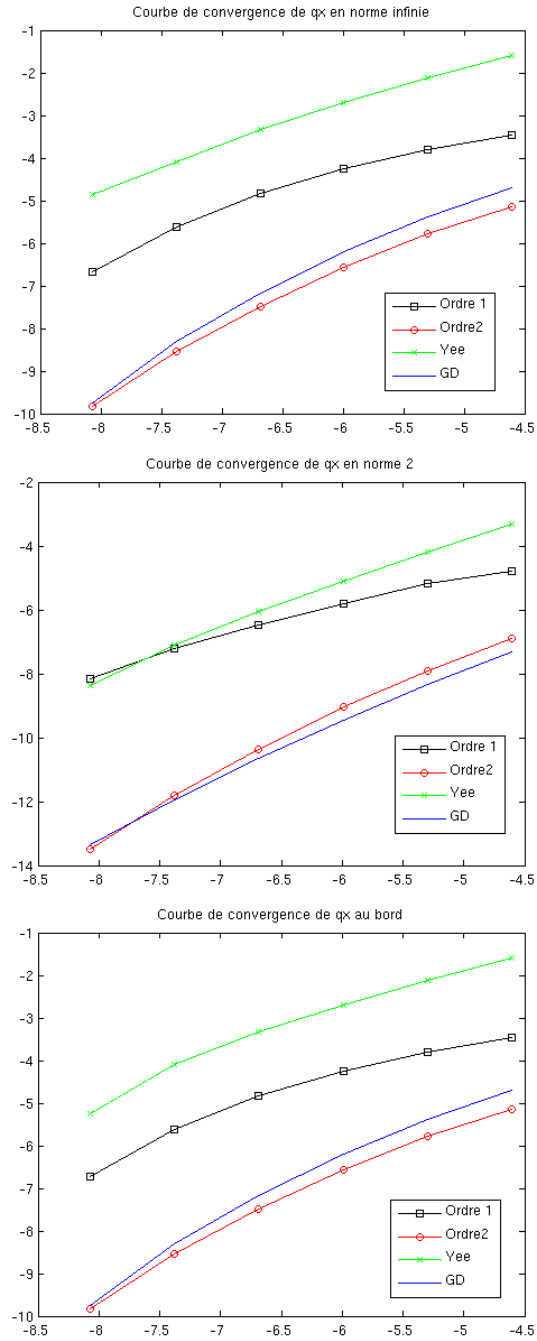


FIGURE 4.15: Courbes de convergence du flux de chaleur

	$\rho$	$u$	$v$	$q_x$	$T$
VF 1	0.87	0.95	1.05	0.91	1.04
VF 2	1.64	1.44	1.61	1.35	1.62
Yee	1.03	0.90	1.20	0.94	1.18
GD	1.28	1.46	1.29	1.44	1.30

	$\rho$	$u$	$v$	$q_x$	$T$
VF 1	0.91	1.52	1.21	0.98	1.06
VF 2	1.97	2.12	1.96	1.89	1.96
Yee	1.66	1.90	1.85	1.44	1.84
GD	1.60	1.50	1.58	1.74	1.64

TABLE 4.2: ordres de convergence en norme  $L^\infty$  et en norme  $L^2$ 

On constate sur les courbes obtenues en norme  $L^2$  que les ordres de convergence sont ceux attendus pour tous les schémas (le schéma Galerkin Discontinu avec un flux upwind étant théoriquement d'ordre  $3/2$ ), sauf le schéma de Yee dont les performances sont un peu dégradées par les conditions aux limites inadaptées. On voit également sur ces courbes que sur les maillages grossiers, où la couche limite cinétique est peu ou pas maillée, le schéma GD donne une erreur moyenne inférieure aux schémas Volumes Finis, conformément aux propriétés attendues.

En revanche, concernant la norme  $L^\infty$ , si les résultats obtenus sur le schéma VF1 sont conformes à ce qui était prévu, on distingue deux comportements sur les schémas VF2 et GD. D'un côté le flux de chaleur et la vitesse horizontale, sur lesquels les deux schémas produisent une erreur quasi identique, nettement inférieure à celle donnée par le schéma VF1, mais avec un ordre de convergence assez bas (situation encore pire pour le schéma de Yee). De l'autre, les courbes obtenues sur la densité, la vitesse verticale et la température, sur lesquels le schéma GD a un ordre proche de 1 tout en restant nettement meilleur que le schéma VF1, et où le schéma VF2, d'ordre plus élevé bien que restant inférieur à 2, montre une chute de l'erreur accélérée sur les maillages fins.

Le calcul du flux de chaleur dans la première cellule du maillage montre que c'est près de la paroi que l'erreur maximale est commise : cette courbe est en effet identique à celle obtenue pour la norme  $L^\infty$ .

En conclusion, on obtient donc les résultats espérés : le schéma Galerkin Discontinu est en moyenne assez nettement meilleur que les schémas Volumes Finis sur les maillages grossiers, où son erreur maximale est équivalente à celle donnée par le schéma d'ordre 2.

#### 4.5.2 Maillages non réguliers

Les schémas VF2 et GD ayant montré des performances comparables sur des maillages réguliers, on les compare maintenant sur des maillages raffinés en proche paroi.

Le raffinement des maillages près des parois est obtenu de la façon suivante, inspirée de [69] : disposant d'un maillage régulier du segment  $[0, 1]$ , dont les points sont notés  $x_i$ ,  $1 \leq i \leq n$ , on applique la transformation

$$\forall 1 \leq i \leq n, \quad \begin{cases} z_i = \exp \left[ \log \left( \frac{\beta + 1}{\beta - 1} \right) (1 - x_i) \right], & \beta > 1, \\ y_i = \frac{1 - z_i}{1 + z_i} \beta + 1, \end{cases}$$

ce qui donne un maillage  $\{y_i, 1 \leq i \leq n\}$  d'autant plus raffiné au bord du domaine de calcul que le paramètre  $\beta$  est proche de 1. Le nombre de cellules étant conservé par la transformation,

le resserrement du maillage près des bords entraîne un relâchement dans le centre du domaine de calcul.

Pour chaque maillage, on compare les résultats avec ceux obtenus sur un maillage régulier de pas  $\Delta x'$  égal à la taille de la plus petite maille du maillage initial (c'est-à-dire de la maille de bord) :

$$\Delta x' = \min(\Delta x_i).$$

La grandeur physique ayant le plus d'intérêt à la paroi est le flux de chaleur : on compare pour différents maillages ce flux dans la première maille, le résultat obtenu sur le maillage régulier fin servant de référence. On choisit  $n = 100$ .

Les cinq maillages utilisés correspondent à des valeurs du paramètre  $\beta$  de  $5.10^{-5}$ ,  $3.10^{-4}$ ,  $5.10^{-4}$ ,  $10^{-3}$  et  $5.10^{-3}$ , correspondant respectivement à des maillages réguliers de 3572, 1731, 1417, 1083 et 588 mailles. Les résultats obtenus sont exposés sur la figure (4.16), sur laquelle on représente le résidu en fonction de la taille de première maille (en échelle logarithmique).

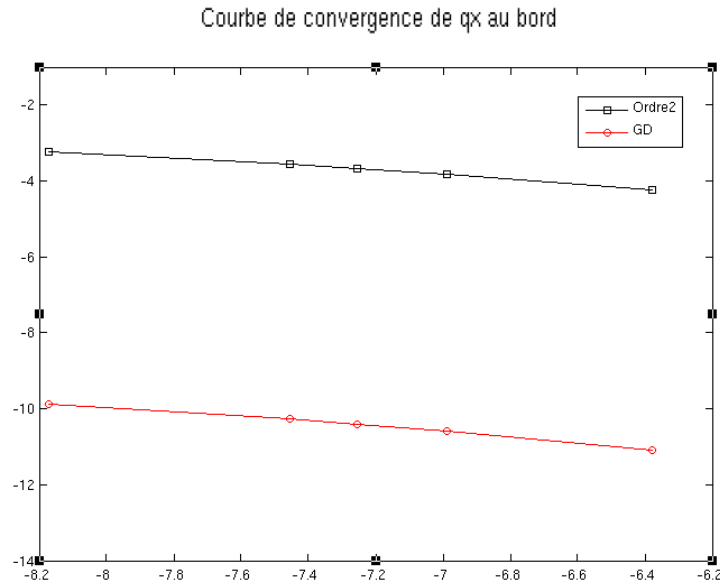


FIGURE 4.16: Flux de chaleur à la paroi en fonction de la taille de première maille (échelle log)

On constate une évolution similaire de l'erreur pour les deux schémas, une légère augmentation avec la concentration des points de calcul sur les bords, ce qui est peut-être dû au déraffinement induit à l'intérieur du domaine, le nombre de points total restant constant. Cependant, l'erreur du schéma GD est bien moindre que celle du schéma VF2.

#### 4.5.3 Vitesse de convergence en temps : analyse des valeurs et espaces propres du système linéaire

La figure (4.10) montre une vitesse de convergence vers l'état stationnaire semblable pour le schéma Galerkin Discontinu et les schémas Volumes Finis d'ordre élevé. On étudie dans cette partie les valeurs et les sous-espaces propres du schéma (4.10) utilisé avec la méthode Volumes Finis ou avec la méthode Galerkin Discontinu afin de comprendre cette évolution.

La seule différence entre les deux schémas étant due à la discrétisation de l'opérateur de transport, nous négligeons dans cette étude l'influence des collisions : on considère l'équation du transport libre avec des conditions aux limites périodiques

$$\forall k \in \mathcal{K}, \quad \begin{cases} \frac{\partial f_k}{\partial t} + v_k \frac{\partial f_k}{\partial x} = 0, \\ f_k|_{\Gamma_1} = f_k|_{\Gamma_2}, \\ f_k(t=0, x) = f_k^0(x) \end{cases} \quad (4.39)$$

où  $\Gamma_1$  et  $\Gamma_2$  sont les deux bords du domaine, et  $f_k^0$  une fonction périodique arbitraire.

On considère un maillage régulier à  $N$  cellules, sur lequel on discrétise le problème (4.39).

Le schéma Volumes Finis s'écrit pour  $v_k > 0$  :

$$\forall 1 \leq i \leq N, \quad f_{i,k}^{n+1} = f_{i,k}^n - \alpha_k \left( f_{i,k}^{n+1} - f_{i-1,k}^{n+1} \right), \quad \alpha_k = v_k \frac{\Delta t}{\Delta x},$$

la condition aux limites n'intervient, compte-tenu de cette écriture, que dans la première cellule du maillage, et donne

$$f_{1,k}^{n+1} = f_{1,k}^n - \alpha_k \left( f_{1,k}^{n+1} - f_{N,k}^{n+1} \right),$$

et on obtient

$$\bar{A}_k^{VF} \mathbf{f}_k^{n+1} = \mathbf{f}_k^n,$$

avec

$$\bar{A}_k^{VF} = \begin{pmatrix} 1 + \alpha_k & & & -\alpha_k \\ -\alpha_k & 1 + \alpha_k & & \\ & \ddots & \ddots & \\ & & -\alpha_k & 1 + \alpha_k \end{pmatrix} \in \mathcal{M}_N(\mathbb{R}), \quad \mathbf{f}_k = (f_{1,k}, \dots, f_{N,k})^T \in \mathbb{R}^N.$$

Le schéma Galerkin Discontinu s'écrit quant à lui selon (4.38), pour  $v_k > 0$  :

$$\forall 1 \leq i \leq N, \quad \mathbf{f}_{i,k}^{n+1} = \mathbf{f}_{i,k}^n + \alpha_k \left[ \begin{pmatrix} 0 & 4 \\ 0 & -2 \end{pmatrix} \mathbf{f}_{i-1,k}^{n+1} + \begin{pmatrix} -3 & -1 \\ 3 & -1 \end{pmatrix} \mathbf{f}_{i,k}^{n+1} \right],$$

la condition aux limites donne

$$\mathbf{f}_{1,k}^{n+1} = \mathbf{f}_{1,k}^n + \alpha_k \left[ \begin{pmatrix} 0 & 4 \\ 0 & -2 \end{pmatrix} \mathbf{f}_{N,k}^{n+1} + \begin{pmatrix} -3 & -1 \\ 3 & -1 \end{pmatrix} \mathbf{f}_{1,k}^{n+1} \right].$$

On obtient donc, sous forme matricielle :

$$\bar{A}_k^{GD} \mathbf{f}_k^{n+1} = \mathbf{f}_k^n,$$

avec

$$\bar{\bar{A}}_k^{GD} = \begin{pmatrix} 1+3\alpha_k & \alpha_k & & & & -4\alpha_k \\ -3\alpha_k & 1+\alpha_k & & & & 2\alpha_k \\ & -4\alpha_k & 1+3\alpha_k & \alpha_k & & \\ & 2\alpha_k & -3\alpha_k & 1+\alpha_k & & \\ & & \ddots & \ddots & \ddots & \ddots \\ & & & & -4\alpha_k & 1+3\alpha_k & \alpha_k \\ & & & & 2\alpha_k & -3\alpha_k & 1+\alpha_k \end{pmatrix} \in \mathcal{M}_{2N}(\mathbb{R}),$$

et

$$\mathbf{f}_k = (\mathbf{f}_{1,k}, \dots, \mathbf{f}_{N,k})^T = (f_{1,k,1}, f_{1,k,2}, \dots, f_{N,k,1}, f_{N,k,2})^T \in \mathbb{R}^{2N}.$$

Les polynômes caractéristiques des matrices  $\bar{\bar{A}}_k^{VF}$  et  $\bar{\bar{A}}_k^{GD}$  sont

$$\begin{cases} |\bar{\bar{A}}_k^{VF} - \lambda \bar{\bar{I}}| = (1 + \alpha_k - \lambda)^N - \alpha_k^N, \\ |\bar{\bar{A}}_k^{GD} - \mu \bar{\bar{I}}| = (Y^2 - 4\alpha_k Y + 6\alpha_k^2)^N - (2\alpha_k Y + 6\alpha_k^2)^N, \quad Y = \mu - 1, \end{cases}$$

ce qui donne les valeurs propres  $\{\lambda_j\}_{0 \leq j \leq N-1}$  de  $\bar{\bar{A}}_k^{VF}$

$$\lambda_j = 1 + \alpha_k \left( 1 - \exp\left(\frac{2ij\pi}{N}\right) \right), \quad 0 \leq j \leq N-1,$$

et  $\{\mu_j^+\}_{0 \leq j \leq N-1}$  et  $\{\mu_j^-\}_{0 \leq j \leq N-1}$  de  $\bar{\bar{A}}_k^{GD}$

$$\begin{cases} \mu_j^+ = 1 + \alpha_k \left( 2 + \exp\left(\frac{2ij\pi}{N}\right) + \left[ \exp\left(\frac{4ij\pi}{N}\right) + 10 \exp\left(\frac{2ij\pi}{N}\right) - 2 \right]^{1/2} \right), & 0 \leq j \leq N-1, \\ \mu_j^- = 1 + \alpha_k \left( 2 + \exp\left(\frac{2ij\pi}{N}\right) - \left[ \exp\left(\frac{4ij\pi}{N}\right) + 10 \exp\left(\frac{2ij\pi}{N}\right) - 2 \right]^{1/2} \right), & 0 \leq j \leq N-1. \end{cases}$$

On trace sur la figure (4.17) les valeurs propres des deux matrices pour  $N = 60$  et  $\alpha_k = 10$  dans le plan complexe.

Les deux matrices  $\bar{\bar{A}}_k^{VF}$  et  $\bar{\bar{A}}_k^{GD}$  ont pour valeur propre 1 (respectivement  $\lambda_0$  et  $\mu_0^-$ ), associée à l'espace des états constants  $\text{Vect}(1, \dots, 1)$ . Les autres valeurs propres sont de module supérieur : les modes propres associés sont donc amortis, et on tend vers un état constant.

La convergence est pilotée par la valeur propre possédant le plus petit module strictement supérieur à 1. Ce sont respectivement pour  $\bar{\bar{A}}_k^{VF}$  et  $\bar{\bar{A}}_k^{GD}$

$$\begin{cases} \lambda_1 = 1 + \alpha_k \left( 1 - \exp\left(\frac{2i\pi}{N}\right) \right), \\ \mu_1^- = 1 + \alpha_k \left( 2 + \exp\left(\frac{2i\pi}{N}\right) - \left[ \exp\left(\frac{4i\pi}{N}\right) + 10 \exp\left(\frac{2i\pi}{N}\right) - 2 \right]^{1/2} \right). \end{cases}$$

Un développement limité au premier ordre donne pour ces deux valeurs

$$\lambda_1, \mu_1^- = 1 - \frac{2i\pi}{N} \alpha_k + \mathcal{O}\left(\frac{1}{N^2}\right),$$

et on a bien un taux de convergence identique, conformément aux observations.

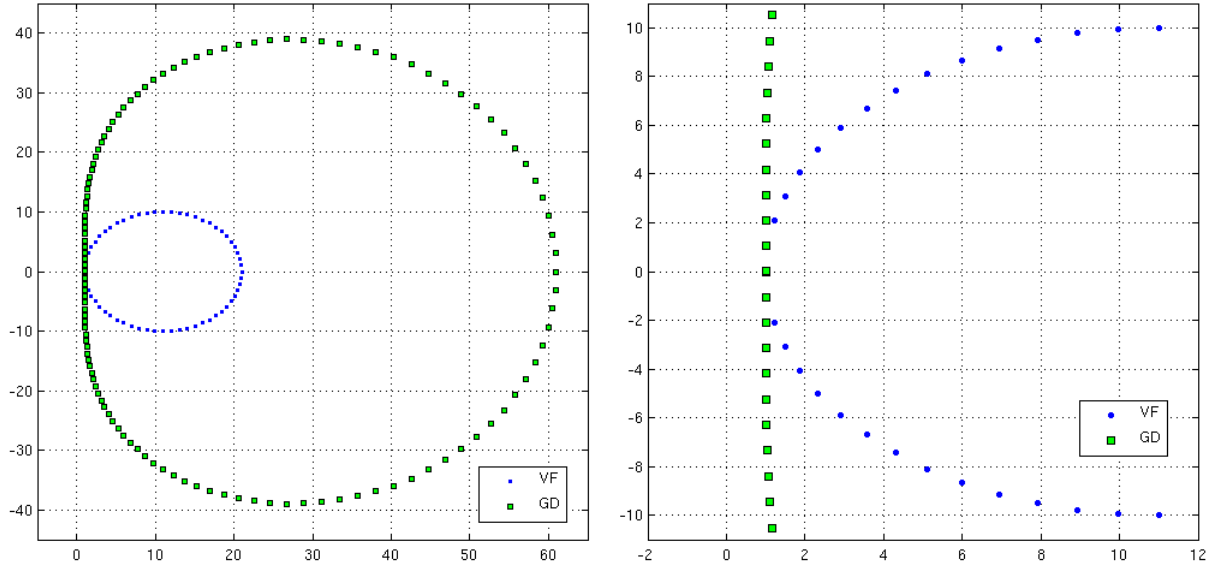


FIGURE 4.17: Valeurs propres des matrices  $\bar{A}_k^{VF}$  et  $\bar{A}_k^{GD}$  pour  $N = 60$  et  $\alpha_k = 10$

On peut naturellement procéder à la même étude pour les vitesses  $v_k < 0$ , ce qui donne les mêmes observations et conclusions.

## 4.6 Conclusion

La comparaison entre les différents schémas Volumes Finis a permis de montrer l'importance d'une discrétisation correcte des conditions aux limites, qui peuvent fortement dégrader les performances d'un schéma si cette discrétisation est inadaptée. L'utilisation d'opérateurs fortement non-linéaires pour augmenter l'ordre de convergence des schémas ne permet pas toujours cette discrétisation appropriée, comme on a pu le voir sur le schéma de Yee ou le schéma d'ordre 4. Le schéma Galerkin Discontinu permet d'éviter cet écueil ; en outre, ses performances globales sont comparables à celles d'un schéma d'ordre 2 sur des maillages réguliers, avec une vitesse de convergence égale. Les tests effectués sur les maillages aérodynamiques montrent une aptitude à être utilisé sur des maillages moins raffinés aux parois que les schémas Volumes Finis d'ordre 2 pour une qualité d'approximation égale, ce qui compense le coût de calcul plus élevé dû à la dimension supérieure de l'espace d'approximation (dans le cas étudié, ce coût est double).

En termes de perspectives, nous pensons donc qu'il serait intéressant de poursuivre cette étude par une comparaison VF/DG en deux dimensions d'espace.

# Conclusion

*Chaque fois que la science avance d'un pas, c'est qu'un imbécile la pousse, sans le faire exprès.*

Émile Zola - *La joie de vivre*

L'utilisation de méthodes numériques déterministes pour la résolution des équations cinétiques de l'aérodynamique en haute altitude se heurte dans les régimes raréfié et transitionnel au problème de la quantité considérable de données à traiter lors du calcul : la fonction de distribution décrite par le modèle cinétique est en effet déterminée dans un espace des phases comptant 3 dimensions de vitesses et usuellement 2 ou 3 dimensions d'espace. La réduction de ces données est un point crucial pour rendre accessibles des calculs d'écoulements présentant des structures complexes (écoulements d'arrière-corps par exemple), des géométries tridimensionnelles, ou des écoulements instationnaires. Dans le même temps, il est évidemment important de conserver une précision élevée des résultats, en particulier dans le calcul des échanges thermiques entre l'air et la paroi du corps de rentrée, échanges qu'il est crucial d'estimer afin d'éviter la désintégration du corps. Il est donc nécessaire de trouver un équilibre entre précision des résultats et coût des calculs.

C'est ce compromis que nous avons tenté d'atteindre dans ce travail : réduire le nombre de points du maillage de l'espace des phases, tout en conservant un niveau de précision des résultats satisfaisant. Nous avons considéré séparément le problème dans l'espace des vitesses et l'espace physique, ce qui a donné lieu à deux approches différentes.

Considérant l'espace des vitesses discrètes, nous avons mis au point une méthode de raffinement local de la grille de vitesses. Cette méthode est fondée sur une approche "physique" du problème : en supposant un faible déséquilibre thermodynamique au sein de l'écoulement, nous avons déterminé pour chaque point de l'espace des vitesses l'allure locale des fonctions de distribution, en particulier la taille de leur support, chacune d'entre elles étant assimilée à la distribution d'équilibre correspondante. Ceci a permis pour une configuration d'écoulement donnée de caractériser les bornes et le pas minimaux que doit posséder la grille de vitesse cartésienne qui est usuellement utilisée pour effectuer le calcul de l'écoulement, puis de construire un maillage localement adapté à la discrétisation de ces fonctions de distribution en utilisant un algorithme récursif. Cette méthode a montré d'excellents résultats sur des géométries 2D planes ou axisymétriques, et surtout sur des géométries 3D, sur lesquels la place mémoire utilisée pour stocker les données et le temps de calcul ont été réduits de façon considérable sans altérer la précision des résultats : nous avons ainsi pu construire des grilles de vitesses localement raffinées comptant jusqu'à 25 fois moins de points que la grille cartésienne régulière minimale nécessaire pour le calcul, avec une différence de moins de 2% dans les flux de chaleur à la paroi et de moins de 3% dans les grandeurs thermodynamiques au sein de l'écoulement.



Dans un second temps, nous avons tenté de réduire également le maillage de l'espace physique. Le compromis à trouver entre la finesse du maillage et la précision des calculs pour un écoulement de gaz raréfié autour d'un corps solide est particulièrement critique dans la zone de proche paroi, où se développe la couche de Knudsen, couche limite très fine qui joue un rôle déterminant dans les échanges thermiques entre le gaz et la paroi. Notre approche fut cette fois plus "mathématique", cherchant à atteindre une réduction de la taille du maillage non plus par la prédiction des propriétés physiques de l'écoulement, mais par une étude des propriétés mathématiques des schémas numériques. En nous inspirant de travaux effectués sur l'équation de transport linéaire, qui ont montré que les schémas Galerkin Discontinu possédaient dans ce cadre une propriété de préservation forte de l'asymptotique, permettant de capturer la solution du problème limite de diffusion sans que le maillage résolve la couche limite cinétique, nous avons étudié l'application de ces schémas à l'équation BGK, dans un cadre adimensionnel linéarisé afin de simplifier l'étude, et cherché à obtenir des propriétés similaires. Après avoir utilisé un développement de Hilbert de la fonction de distribution des particules en fonction du nombre de Knudsen, quantifiant la raréfaction du gaz, nous avons montré que les schémas Galerkin Discontinu préservent fortement la structure de maxwellienne linéarisée des distributions d'équilibre puis que, moyennant dans le cas général une hypothèse d'uniformité sur les perturbations de vitesse et de température à la paroi, les problèmes limites de Laplace et de Boussinesq donnant respectivement la température et la densité à l'ordre 0 du développement sont préservés avec leurs conditions aux limites d'adhérence. En revanche, si la condition d'adhérence sur la vitesse au même ordre est également préservée, le problème de Stokes sous contraintes liant cette vitesse et la pression à l'ordre 1 est mal posé car ses deux inconnues sont approchées par des fonctions polynomiales du même degré, ce que nous n'avons pas la latitude de corriger en introduisant des termes de pénalisation dans la forme faible du problème.

Chacune de ces deux voies ouvre des perspectives. Concernant le raffinement local des grilles de vitesses discrètes, une reconstruction de la grille au cours de la convergence vers l'état stationnaire est envisagée, en utilisant les grandeurs thermodynamiques de l'itération courante pour le calcul de la fonction "couleur" servant de base à la construction d'une nouvelle grille, puis en projetant les données sur cette grille reconstruite. Il est également envisagé d'adapter la construction de la grille AMR au modèle ES-BGK, récemment implanté dans le code d'aérodynamique raréfiée du CEA-CESTA : cela nécessitera l'adaptation du raffinement à l'anisotropie des distributions d'équilibre gaussiennes utilisées dans ce modèle.

L'étude des schémas Galerkin Discontinu appliqués à l'équation BGK peut être étendue. Dans un cadre linéarisé, les calculs effectués dans ce travail jusqu'à l'ordre 0 du développement de Hilbert peuvent être poursuivis à l'ordre suivant. Cependant, on ne peut espérer là encore que des résultats partiels : on doit retrouver à cet ordre une autre équation de Stokes discrète, qui sera *a priori* aussi mal posée que celle obtenue à l'ordre 0. Les premiers calculs effectués dans ce sens laissent de plus entrevoir des calculs longs et complexes. Une autre voie d'exploration possible est le passage au cadre non linéarisé, afin de rechercher l'asymptotique Navier-Stokes.

## Annexe A

# Overview of the 3D kinetic code

### A.1 The linearized implicit scheme

Our code is an extension of the code presented in [52] to 3D polyatomic flows. It is based on the following reduced BGK model

$$\begin{aligned}\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f &= \frac{1}{\tau} (M(U) - f) \\ \partial_t g + \mathbf{v} \cdot \nabla_{\mathbf{x}} g &= \frac{1}{\tau} \left( \frac{\delta}{2} RT M(U) - g \right),\end{aligned}$$

where  $U = \langle mf + e^{(5)}g \rangle = (\rho, \rho \mathbf{u}, E = \frac{1}{2}\rho |\mathbf{u}|^2 + \frac{3+\delta}{2}\rho RT)$  is the vector of macroscopic mass, momentum, and energy density. Here, we use the standard notation  $\langle . \rangle = \int_{\mathbb{R}^3} . d\mathbf{v}$  for any vector valued function of  $\mathbf{v}$ ,  $m(v) = (1, \mathbf{v}, \frac{1}{2}|\mathbf{v}|^2)$  is the vector of collisional invariants, and  $e^{(5)} = (0, 0, 0, 0, 1)$ . Moreover,  $M(U)$  is the standard Maxwellian distribution defined through the density, velocity, and temperature corresponding to the vector  $U$  above (see (1.2)).

This model comes from the reduction of a BGK model for the full distribution function  $F(t, \mathbf{x}, \mathbf{v}, I)$ , where  $I$  is the internal energy variable, and  $\delta$  is the number of internal degrees of freedom (see [39] for the first use of this technique and [3] for an application to polyatomic gases). Consequently, it accounts for any number of internal degrees of freedom. For instance, a diatomic gas can be described with  $\delta = 2$ .

This model is first discretized with respect to the velocity variable. We follow the approximation of [52] and its extension to polyatomic gases [33]. We assume we have a velocity grid  $\{\mathbf{v}_q, q = 0 : q_{max}\}$  (like the grids described in this paper). The continuous distributions  $f$  and  $g$  are then replaced by their approximations at each point  $\mathbf{v}_q$ , and we get the following discrete velocity BGK system

$$\begin{aligned}\partial_t f_q + \mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q &= \frac{1}{\tau} (M_q(U) - f_q) \\ \partial_t g_q + \mathbf{v}_q \cdot \nabla_{\mathbf{x}} g_q &= \frac{1}{\tau} (N_q(U) - g_q),\end{aligned}$$

where  $(M_q(U), N_q(U))$  is an approximation of  $(M(U), \frac{\delta}{2}RTM(U))$  that has to be defined. As explained in this paper, we assume we have quadrature weights  $\{\omega_q\}$  corresponding to our discrete velocity grid, so that the moment vector of the discrete distributions is

$$U = \sum_{q=0}^{q_{max}} (m(\mathbf{v}_q) f_q + e^{(5)} g_q) \omega_q.$$

As proposed in [33], the discrete equilibrium  $(M_q(U), N_q(U))$  is constructed so that it has the same moments as  $(f, g)$ , that is to say

$$\sum_{q=0}^{q_{max}} (m(\mathbf{v}_q) M_q(U) + e^{(5)} N_q(U)) \omega_q = U. \quad (\text{A.1})$$

In our code,  $(M_q(U), N_q(U))$  is determined through the entropic variable  $\alpha$  such that  $M_q(U) = \exp(\alpha \cdot m(\mathbf{v}_q))$  and  $N_q(U) = \frac{\delta}{2} \frac{1}{-\alpha_5} \exp(\alpha \cdot m(\mathbf{v}_q))$ , by solving (A.1) by a Newton algorithm. We mention that the computational cost of this algorithm can be significantly reduced by using a nice idea due to Titarev [67]. This optimization will be used in a future version of our code.

The discrete velocity BGK system is then discretized by a finite volume scheme on a multiblock curvilinear 3D mesh of hexahedral cells  $\Omega_{ijk}$ , with indices  $i, j, k = 1$  to  $i_{max}, j_{max}, k_{max}$ , respectively. Denoting by  $f_{ijkq}^n$  an approximation of the average of  $f$  at time  $t_n$  on a cell  $\Omega_{ijk}$  at the discrete velocity  $\mathbf{v}_q$ , our scheme reads, in its implicit version,

$$\begin{aligned} \frac{f_{ijkq}^{n+1} - f_{ijkq}^n}{\Delta t} + (\mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q^{n+1})_{ijk} &= \frac{1}{\tau_{ijk}^{n+1}} (M_q(U_{ijk}^{n+1}) - f_{ijkq}^{n+1}), \\ \frac{g_{ijkq}^{n+1} - g_{ijkq}^n}{\Delta t} + (\mathbf{v}_q \cdot \nabla_{\mathbf{x}} g_q^{n+1})_{ijk} &= \frac{1}{\tau_{ijk}^{n+1}} (N_q(U_{ijk}^{n+1}) - g_{ijkq}^{n+1}), \end{aligned}$$

where

$$U_{ijk}^{n+1} = \sum_{q=0}^{q_{max}} (m(\mathbf{v}_q) f_{ijkq}^{n+1} + e^{(5)} g_{ijkq}^{n+1}) \omega_q.$$

The discrete divergence  $(\mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q)_{ijk}^{n+1}$  is given by the following second order upwind approximation (with the Yee limiter [73]) :

$$(\mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q^{n+1})_{ijk} = \frac{1}{|\Omega_{i,j,k}|} \left[ \left( \Phi_{i+\frac{1}{2},j,k} - \Phi_{i-\frac{1}{2},j,k} \right) + \left( \Phi_{i,j+\frac{1}{2},k} - \Phi_{i,j-\frac{1}{2},k} \right) + \left( \Phi_{i,j,k+\frac{1}{2}} - \Phi_{i,j,k-\frac{1}{2}} \right) \right], \quad (\text{A.2})$$

where

$$\Phi_{i+\frac{1}{2},j,k} = (\mathbf{v}_q \cdot \nu_{i+\frac{1}{2},j,k})^+ f_{ijkq}^{n+1} + (\mathbf{v}_q \cdot \nu_{i+\frac{1}{2},j,k})^- f_{i+1,j,k,q}^{n+1} + |\mathbf{v}_q \cdot \nu_{i+\frac{1}{2},j,k}| \minmod(\Delta_{i-\frac{1}{2}}^{n+1}, \Delta_{i+\frac{1}{2}}^{n+1}, \Delta_{i+\frac{3}{2}}^{n+1})$$

is the second order numerical flux across the face between  $\Omega_{i,j,k}$  and  $\Omega_{i+1,j,k}$ , and  $\nu_{i+\frac{1}{2},j,k}$  is the normal vector to this face directed from  $\Omega_{i,j,k}$  to  $\Omega_{i+1,j,k}$  while its norm is equal to the area of the face. In the minmod limiter function, we use the notation  $\Delta_{i+\frac{1}{2}}^{n+1} = f_{i+1,j,k,q}^{n+1} - f_{ijkq}^{n+1}$ . Finally, we use the standard notation  $a^\pm = (a \pm |a|)/2$  for every number  $a$ . The numerical fluxes across the other faces are defined accordingly. For the sequel, it is useful to denote by  $(\mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q^{n+1})_{ijk}^{1st\ order}$  the corresponding linear first order upwind discretization (that is to say, defined by (A.2) where the minmod term is set to 0).

The advantage of the time implicit approximation is that it ensures unconditional stability, which allows us to take large time steps, and hence to get rapid convergence to the steady state. Of course, this scheme is too expensive, since it requires to solve a non linear system at each time iteration. Therefore, we rather use the following linearization. First, the Maxwellian is linearized by a first order Taylor expansion :

$$M_q(U_{ijk}^{n+1}) = M_q(U_{ijk}^n + (U_{ijk}^{n+1} - U_{ijk}^n)) \approx M_q(U_{ijk}^n) + \partial_U M_q(U_{ijk}^n)(U_{ijk}^{n+1} - U_{ijk}^n),$$

and the same for  $N_q(U_{ijk}^{n+1})$  (see section A.4 for explicit expressions of the Jacobian matrices). Then the discrete divergence, which is not differentiable due to the limiter, is linearized by using the corresponding first order upwind approximation. This is sometimes called a “frozen coefficient technique”, which gives

$$\begin{aligned} (\mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q)_{ijk}^{n+1} &= \left( \mathbf{v}_q \cdot \nabla_{\mathbf{x}} (f_q^n + (f_q^{n+1} - f_q^n)) \right)_{ijk} \\ &\approx \left( \mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q^n \right)_{ijk} + \left( \mathbf{v}_q \cdot \nabla_{\mathbf{x}} (f_q^{n+1} - f_q^n) \right)_{ijk}^{1st\ order} \end{aligned}$$

Then, denoting by  $\delta f_{i,j,k,q} = f_{ijkq}^{n+1} - f_{ijkq}^n$  (same notation for  $g$ ), and by  $\delta U_{i,j,k}$  the moments of  $(\delta f_{i,j,k}, \delta g_{i,j,k})$ , our scheme reads in the following form :

$$\frac{\delta f_{i,j,k,q}}{\Delta t} + (\mathbf{v}_q \cdot \nabla_{\mathbf{x}} \delta f_q)_{ijk}^{1st\ order} - \frac{1}{\tau_{i,j,k}^n} \partial_U M_q(U_{ijk}^n)(\delta U_{i,j,k}) = RHS f_{i,j,k}^n \quad (\text{A.3})$$

$$\frac{\delta g_{i,j,k,q}}{\Delta t} + (\mathbf{v}_q \cdot \nabla_{\mathbf{x}} \delta g_q)_{ijk}^{1st\ order} - \frac{1}{\tau_{i,j,k}^n} \partial_U N_q(U_{ijk}^n)(\delta U_{i,j,k}) = RHS g_{i,j,k}^n, \quad (\text{A.4})$$

where the right-hand sides are given by

$$\begin{aligned} RHS f_{i,j,k}^n &= - \left( \mathbf{v}_q \cdot \nabla_{\mathbf{x}} f_q^n \right)_{ijk} + \frac{1}{\tau_{ijk}^n} (M_q(U_{ijk}^n) - f_{ijkq}^n) \\ RHS g_{i,j,k}^n &= - \left( \mathbf{v}_q \cdot \nabla_{\mathbf{x}} g_q^n \right)_{ijk} + \frac{1}{\tau_{ijk}^n} (N_q(U_{ijk}^n) - g_{ijkq}^n). \end{aligned}$$

If our scheme converges to steady state, then the right-hand side is zero, and we get a second order discrete steady solution.

## A.2 Numerical boundary conditions

Numerically, the boundary conditions are implemented by the standard ghost cell technique, which is used as follows. When an index  $(i, j, k)$  corresponds to a cell located at the boundaries of the domain, there appear unknown values in the numerical fluxes like  $f_{i,0,k,q}^n$  and  $f_{i,j_{max}+1,k,q}^n$ , for the cells  $(i, 1, k)$  and  $(i, j_{max}, k)$  for instance (see (A.2)). Corresponding cells  $(i, 0, k)$ ,  $(i, j_{max} + 1, k)$ , etc. are called ghost-cells. These values are classically defined according to the boundary conditions that are specified for the problem. Here we use several kinds of boundary conditions : solid wall interactions, inflow and outflow boundary conditions at artificial boundaries, as well as symmetry boundary conditions along symmetry planes and symmetry axes.

For the diffuse reflection, the incident molecules in a boundary cell of index  $(i, 1, k)$  are supposed to be re-emitted by the wall from a ghost cell of index  $(i, 0, k)$ . This cell is the mirror

cell of  $(i, 1, k)$  with respect to the wall. The diffuse reflection (1.5)–(1.6) is then modeled by

$$(f_{i,0,k,q}^n, g_{i,0,k,q}^n) = \sigma_{i,1,k} (M_q^{wall}, N_q^{wall}), \quad \mathbf{v}_q \cdot \mathbf{n}_{i,1,k} > 0, \quad (\text{A.5})$$

where  $\sigma_{i,1,k}$  is determined so as to avoid a mass flux across the wall, that is to say between cells  $(i, 0, k)$  and  $(i, 1, k)$ . In this relation,  $(M_q^{wall}, N_q^{wall})$  is a discrete conservative approximation of the wall Maxwellians  $\left(1, \frac{\delta}{2} RT_{wall}\right) \frac{1}{(2\pi RT_{wall})^{3/2}} \exp\left(-\frac{|\mathbf{v}|^2}{2RT_{wall}}\right)$ . Relation (1.6) gives

$$\sigma_{i,1,k} = - \left( \sum_{\mathbf{v}_q \cdot \mathbf{n}_{i,1,k} < 0} \mathbf{v}_q \cdot \mathbf{n}_{i,1,k} f_{i,1,k,q}^n \omega_q \right) \left( \sum_{\mathbf{v}_q \cdot \mathbf{n}_{i,1,k} > 0} \mathbf{v}_q \cdot \mathbf{n}_{i,1,k} M_q^{wall} \omega_q \right)^{-1}.$$

For the inflow boundary condition, for instance at a boundary cell  $(i, j_{max} + 1, k)$ , we simply set the ghost cell value to the upstream Maxwellian distributions  $(M_q^{upstream}, N_q^{upstream})$  (defined through the upstream values of  $\rho$ ,  $u$ , and  $T$ ) :

$$(f_{i,j_{max}+1,k,q}^n, g_{i,j_{max}+1,k,q}^n) = (M_q^{upstream}, N_q^{upstream}).$$

For the outflow boundary condition, we set the ghost cell value to the value of the corresponding boundary cell :

$$(f_{i,j_{max}+1,k,q}^n, g_{i,j_{max}+1,k,q}^n) = (f_{i,j_{max},k,q}^n, g_{i,j_{max},k,q}^n).$$

Finally, for a cell  $(1, j, k)$  in a symmetry plane (for instance the plane  $(0, y, z)$ ) we use the symmetry of the distribution functions to set

$$(f_{0,j,k,q}^n, g_{0,j,k,q}^n) = (f_{1,j,k,q'}^n, g_{1,j,k,q'}^n),$$

where  $q'$  is such that  $\mathbf{v}_{q'}$  is the symmetric of  $\mathbf{v}_q$  with respect to the symmetry plane. In this case, our AMR velocity grid is constructed such that it is also symmetric with respect to this plane, which gives  $\mathbf{v}_{q'} = (-v_{x,q}, v_{y,q}, v_{z,q})$ . This grid is obtained in two steps : first, a part on one side of the symmetry plane is obtained by using our algorithms (support function and AMR grid generation), and then this part is symmetrized to obtain the part on the other side.

For the second order numerical flux, we also need the values of a second layer of ghost cells with indices like  $(-1, j, k)$  or  $(i_{max} + 2, j, k)$ , etc. For these values, we simply copy the value of the corresponding ghost cell in the first layer, that is to say  $f_{-1,j,k,q}^n = f_{0,j,k,q}^n$  for instance. This treatment makes the accuracy of our scheme reduce to first order at the boundary (since this makes the flux limiters equal to 0). It may be more relevant to use extrapolation techniques, which is the subject of a work in progress. Only the boundary condition on a symmetry plane is treated differently : here, we use the symmetry of the distributions to set

$$(f_{-1,j,k,q}^n, g_{-1,j,k,q}^n) = (f_{2,j,k,q'}^n, g_{2,j,k,q'}^n),$$

where  $v_{q'}$  has been defined above.

### A.3 Linear solver

At each time iteration, our linearized implicit scheme requires to solve the large linear system (A.3–A.4). It is therefore interesting to write it in the following matrix form :

$$\left(\frac{I}{\Delta t} + T + R^n\right)\delta F = RHS^n,$$

where  $\delta F = (\delta f_{i,j,k,q}, \delta g_{i,j,k,q})$  is a large vector that stores all the unknowns,  $I$  is the unit matrix,  $T$  is a matrix such that

$$(T\delta F)_{i,j,k,q} = \left( \left( \mathbf{v}_q \cdot \nabla_{\mathbf{x}} \delta f_q \right)_{ijk}^{1st\ order}, \left( \mathbf{v}_q \cdot \nabla_{\mathbf{x}} \delta g_q \right)_{ijk}^{1st\ order} \right), \quad (\text{A.6})$$

$R^n$  is the relaxation matrix such that

$$(R^n \delta F)_{i,j,k,q} = \left( -\frac{1}{\tau_{i,j,k}^n} \partial_U M_q(U_{ijk}^n) (\delta U_{i,j,k}), -\frac{1}{\tau_{i,j,k}^n} \partial_U N_q(U_{ijk}^n) (\delta U_{i,j,k}) \right), \quad (\text{A.7})$$

and  $RHS^n = (RHS f_{i,j,k}^n, RHS g_{i,j,k}^n)$ .

For simplicity, we use explicit boundary conditions, which means  $\delta F_{i,j,k,q} = 0$  in ghost cells. This implies that  $T$  has a simple block structure, which is used in the following.

The algorithm used in our code is based on a coupling between the iterative Jacobi and Gauss-Seidel methods. First, the relaxation matrix  $R^n$  is splitted into its diagonal  $\Delta^n$  and its off-diagonal  $-E^n$ , so that we get the following system (this the Jacobi iteration) :

$$\left(\frac{I}{\Delta t} + T + \Delta^n\right)\delta F = RHS^n + E^n \delta F. \quad (\text{A.8})$$

Note that  $E^n$  is very sparse : the product  $E^n \delta F$  is local in space and can be written  $[E_{i,j,k}^n \delta F_{i,j,k}]_q$  (see section A.4). The left-hand side of this system has a three level tridiagonal block structure. We solve it by using a line Gauss-Seidel iteration which is described below.

With curvilinear grids, in many cases, and in particular for re-entry problems, the flow is aligned with a mesh (generally aligned with solid boundaries), which means that the largest variations occur along the orthogonal direction, say the direction of index  $i$ , for instance. The idea is to use an “exact” inversion of  $T$  along this direction. This is done with a sweeping strategy sometimes called “Gauss-Seidel line iteration”. First, we rewrite system (A.8) pointwise, as follows :

$$\begin{aligned} & \left( \frac{1}{\Delta t} + A + B + C + \Delta_{i,j,k,q}^n \right) \delta F_{i,j,k,q} + A^- \delta F_{i+1,j,k,q} + A^+ \delta F_{i-1,j,k,q} \\ & + B^- \delta F_{i,j+1,k,q} + B^+ \delta F_{i,j-1,k,q} + C^- \delta F_{i,j,k+1,q} + C^+ \delta F_{i,j,k-1,q} = RHS_{i,j,k,q}^n + [E_{i,j,k}^n \delta F_{i,j,k}]_q, \end{aligned}$$

where coefficients  $A, B, C, A^\pm, B^\pm, C^\pm$  are standard notations for line-Gauss-Seidel method and can be easily derived from (A.8), (A.6), and (A.2). Then the linear solver reads as shown in Algorithm 1.

**Algorithm 1** Jacobi/line-Gauss-Seidel algorithm

---

```

 $\delta F^{(0)} = 0$  % initialization
for  $p$  from 0 to  $P$  do % iterations of the solver
  for  $q$  from 0 to  $q_{max}$  do % loop over the discrete velocities (Jacobi loop)

    % one sweep in  $j$  direction (forward substitution)
    for  $k$  from 1 to  $k_{max}$  do
      for  $j$  from 1 to  $j_{max}$  do
        solve exactly the tridiagonal system

          
$$\left( \frac{1}{\Delta t} + A + B + \Delta_{i,j,k,q}^n \right) \delta F_{i,j,k,q}^{(p+\frac{1}{2})} + A^- \delta F_{i+1,j,k,q}^{(p+\frac{1}{2})} + A^+ \delta F_{i-1,j,k,q}^{(p+\frac{1}{2})}$$

          
$$= RHS_{i,j,k,q}^n + [E_{i,j,k}^n \delta F_{i,j,k}^{(p)}]_q - B^+ \delta F_{i,j-1,k,q}^{(p+\frac{1}{2})} - B^- \delta F_{i,j+1,k,q}^{(p)}$$


        end for
      end for

    % one sweep in  $k$  direction (forward substitution)
    for  $j$  from 1 to  $j_{max}$  do
      for  $k$  from 1 to  $k_{max}$  do
        solve exactly the tridiagonal system

          
$$\left( \frac{1}{\Delta t} + A + C + \Delta_{i,j,k,q}^n \right) \delta F_{i,j,k,q}^{(p+1)} + A^- \delta F_{i+1,j,k,q}^{(p+1)} + A^+ \delta F_{i-1,j,k,q}^{(p+1)}$$

          
$$= RHS_{i,j,k,q}^n + [E_{i,j,k}^n \delta F_{i,j,k}^{(p)}]_q - C^+ \delta F_{i,j,k-1,q}^{(p+1)} - C^- \delta F_{i,j,k+1,q}^{(p+\frac{1}{2})}$$


        end for
      end for
    end for

    compute the moments of  $\delta F^{(p+1)}$  % for the action of  $\Delta_{i,j,k,q}^n$  and  $E_{i,j,k}^n$  onto  $\delta F^{(p+1)}$ 

  end for
  set  $\delta F = \delta F^{(P+1)}$ 

```

---

- Remark A.1.*
1. We observe that we get a fast convergence to steady state by using a few step of this linear solver (say  $P = 2$  or  $3$ ). This is due to the “exact” inversion along the direction of largest variation of the flow.
  2. In practice, we add one back substitution step right after each forward substitution step.
  3. This solver is a straightforward extension of the linear solver used in [52]. More sophisticated versions of this solver could be tested.

## A.4 Jacobian matrices and the relaxation matrix $R^n$

Elementary calculus gives the following formula :

$$\begin{aligned}\partial_U M_q(U) &= M_q(U) m(\mathbf{v}_q) A(U)^{-1} \\ \partial_U N_q(U) &= N_q(U) \left( m(\mathbf{v}_q) - \frac{1}{\alpha_5(U)} e^{(5)} \right) A(U)^{-1},\end{aligned}$$

where  $A(U)$  is the following  $5 \times 5$  matrix

$$A(U) = \sum_{q=0}^{q_{max}} \left( m(\mathbf{v}_q)^T m(\mathbf{v}_q) M_q(U) + e^{(5)T} \left( m(\mathbf{v}_q) - \frac{1}{\alpha_5(U)} e^{(5)} \right) N_q(U) \right) \omega_q.$$

Consequently, by using the definition (A.7) of the relaxation matrix  $R^n$ , a careful algebra shows that its diagonal elements  $\Delta_{i,j,k,q}^n$  can be separated into the following two blocks ( $\Delta_{i,j,k,q}^{n,f}, \Delta_{i,j,k,q}^{n,g}$ ) :

$$\begin{aligned}\Delta_{i,j,k,q}^{n,f} &= \frac{1}{\tau_{i,j,k}^n} \left( M_q[U_{ijk}^n] m(\mathbf{v}_q) A(U_{ijk}^n)^{-1} m(\mathbf{v}_q)^T \omega_q - 1 \right), \\ \Delta_{i,j,k,q}^{n,g} &= \frac{1}{\tau_{i,j,k}^n} \left( N_q[U_{ijk}^n] \left( m(\mathbf{v}_q) - \frac{1}{\alpha_5(U_{ijk}^n)} e^{(5)} \right) A(U_{ijk}^n)^{-1} e^{(5)T} \omega_q - 1 \right).\end{aligned}$$

Therefore, the product of the off-diagonal part  $-E^n$  of  $R^n$  with  $\delta F$  (as used in algorithm 1) is simply

$$\begin{aligned}-[E_{i,j,k}^n \delta F_{i,j,k}]_q &= \left( -\frac{1}{\tau_{i,j,k}^n} \partial_U M_q(U_{ijk}^n) (\delta U_{i,j,k}) - \Delta_{i,j,k,q}^{n,f} \delta f_{i,j,k,q} \right. \\ &\quad \left. , -\frac{1}{\tau_{i,j,k}^n} \partial_U N_q(U_{ijk}^n) (\delta U_{i,j,k}) - \Delta_{i,j,k,q}^{n,g} \delta g_{i,j,k,q} \right).\end{aligned}$$





## Annexe B

# Équilibre linéarisé et opérateur $\Pi$

### B.1 Linéarisation de la maxwellienne

On linéarise  $M[\rho + \rho', \mathbf{u} + \mathbf{u}', T + T']$  au 1<sup>er</sup> ordre en supposant  $\rho' \ll \rho$ ,  $\mathbf{u}' \ll \mathbf{u}$ ,  $T' \ll T$  :

$$M' = M[\rho + \rho', \mathbf{u} + \mathbf{u}', T + T'] = \underbrace{\frac{\rho + \rho'}{(2\pi(T + T'))^{\frac{3}{2}}}}_{(a)} \underbrace{\exp\left(-\frac{|\mathbf{u} + \mathbf{u}' - \mathbf{v}|^2}{2(T + T')}\right)}_{(b)}.$$

On développe chaque terme à l'ordre 1 :

$$\begin{aligned} (a) &= \frac{\rho + \rho'}{(2\pi(T + T'))^{\frac{3}{2}}} = \frac{\rho}{(2\pi T)^{\frac{3}{2}}} \left(1 + \frac{\rho'}{\rho}\right) \left(1 + \frac{T'}{T}\right)^{-\frac{3}{2}} \\ &= \frac{\rho}{(2\pi T)^{\frac{3}{2}}} \left(1 + \frac{\rho'}{\rho}\right) \left(1 - \frac{3}{2} \frac{T'}{T} + \mathcal{O}(T'^2)\right) \\ &= \frac{\rho}{(2\pi T)^{\frac{3}{2}}} \left(1 + \frac{\rho'}{\rho} - \frac{3}{2} \frac{T'}{T} + \mathcal{O}(\rho'^2, T'^2)\right), \end{aligned}$$

$$\begin{aligned} (b) &= \exp\left(-\frac{|\mathbf{u} + \mathbf{u}' - \mathbf{v}|^2}{2(T + T')}\right) = \exp\left(-\frac{|\mathbf{u} + \mathbf{u}' - \mathbf{v}|^2}{2T} \left(1 + \frac{T'}{T}\right)^{-1}\right) \\ &= \exp\left(-\frac{|\mathbf{u} + \mathbf{u}' - \mathbf{v}|^2}{2T} \left(1 - \frac{T'}{T} + \mathcal{O}(T'^2)\right)\right) \\ &= \exp\left(|\mathbf{u} + \mathbf{u}' - \mathbf{v}|^2 \left(\frac{-1}{2T} + \frac{T'}{2T^2} + \mathcal{O}(T'^2)\right)\right) \\ &= \exp\left(\left(|\mathbf{u} - \mathbf{v}|^2 + 2(\mathbf{u} - \mathbf{v}) \cdot \mathbf{u}' + \mathcal{O}(\mathbf{u}'^2)\right) \left(\frac{-1}{2T} + \frac{T'}{2T^2} + \mathcal{O}(T'^2)\right)\right) \\ &= \exp\left(\frac{-|\mathbf{u} - \mathbf{v}|^2}{2T} + \frac{|\mathbf{u} - \mathbf{v}|^2 T'}{2T^2} - \frac{(\mathbf{u} - \mathbf{v}) \cdot \mathbf{u}'}{T} + \mathcal{O}(\mathbf{u}'^2, T'^2)\right) \\ &= \exp\left(\frac{-|\mathbf{u} - \mathbf{v}|^2}{2T}\right) \exp\left(\frac{|\mathbf{u} - \mathbf{v}|^2 T'}{2T^2} - \frac{(\mathbf{u} - \mathbf{v}) \cdot \mathbf{u}'}{T} + \mathcal{O}(\mathbf{u}'^2, T'^2)\right) \end{aligned}$$

$$= \exp\left(\frac{-|\mathbf{u} - \mathbf{v}|^2}{2T}\right) \left(1 - \frac{(\mathbf{u} - \mathbf{v}) \cdot \mathbf{u}'}{T} + \frac{|\mathbf{u} - \mathbf{v}|^2}{2T^2} T' + \mathcal{O}(\mathbf{u}'^2, T'^2)\right).$$

On obtient donc :

$$\begin{aligned} M' &= \underbrace{\frac{\rho}{(2\pi T)^{\frac{3}{2}}} \exp\left(\frac{-|\mathbf{u} - \mathbf{v}|^2}{2T}\right)}_{= M[\rho, \mathbf{u}, T]} \left(1 + \frac{\rho'}{\rho} - \frac{3}{2} \frac{T'}{T} + \mathcal{O}(\rho'^2, T'^2)\right) \\ &\quad \times \left(1 - \frac{(\mathbf{u} - \mathbf{v}) \cdot \mathbf{u}'}{T} + \frac{|\mathbf{u} - \mathbf{v}|^2}{2T^2} T' + \mathcal{O}(\mathbf{u}'^2, T'^2)\right) \\ &= M[\rho, \mathbf{u}, T] \left(1 - \frac{(\mathbf{u} - \mathbf{v}) \cdot \mathbf{u}'}{T} + \frac{|\mathbf{u} - \mathbf{v}|^2}{2T^2} T' + \frac{\rho'}{\rho} - \frac{3}{2} \frac{T'}{T}\right) + \mathcal{O}(\rho'^2, \mathbf{u}'^2, T'^2), \end{aligned}$$

soit :

$$\begin{aligned} M[\rho + \rho', \mathbf{u} + \mathbf{u}', T + T'] &= M[\rho, \mathbf{u}, T] \left[1 + \frac{\rho'}{\rho} + \frac{\mathbf{v} - \mathbf{u}}{T} \cdot \mathbf{u}' + \left(\frac{|\mathbf{v} - \mathbf{u}|^2}{2T^2} - \frac{3}{2T}\right) T'\right] \\ &\quad + \mathcal{O}(\rho'^2, \mathbf{u}'^2, T'^2). \end{aligned}$$

## B.2 Moments de l'opérateur $\Pi(\rho', \mathbf{u}', T')$

L'équilibre linéarisé s'écrit :

$$\begin{aligned} \Pi(\rho', \mathbf{u}', T') &= \left[\rho' + \mathbf{u}' \cdot \mathbf{v} + \left(\frac{|\mathbf{v}|^2}{2} - \frac{3}{2}\right) T'\right] M[1, \mathbf{0}, 1] \\ &= \left[\rho' - \frac{3}{2} T' + \mathbf{u}' \cdot \mathbf{v} + \left(\frac{|\mathbf{v}|^2}{2}\right) T'\right] M[1, \mathbf{0}, 1]. \end{aligned}$$

On calcule  $\langle \mathbf{m}(\mathbf{v}) \Pi(\rho', \mathbf{u}', T') \rangle = \langle (1, \mathbf{v}, |\mathbf{v}|^2)^T \Pi(\rho', \mathbf{u}', T') \rangle$  :

$$\begin{aligned} \bullet \int_{\mathbb{R}^3} \Pi(\rho', \mathbf{u}', T') d\mathbf{v} &= \int_{\mathbb{R}^3} \left[\rho' - \frac{3}{2} T' + \mathbf{u}' \cdot \mathbf{v} + \left(\frac{|\mathbf{v}|^2}{2}\right) T'\right] M[1, \mathbf{0}, 1] d\mathbf{v} \\ &= \frac{1}{(2\pi)^{\frac{3}{2}}} \left[ \left(\rho' - \frac{3}{2} T'\right) \underbrace{\int_{\mathbb{R}^3} \exp\left(\frac{-|\mathbf{v}|^2}{2}\right) d\mathbf{v}}_{= (2\pi)^{\frac{3}{2}}} + \sum_i u'_i \underbrace{\int_{\mathbb{R}^3} v_i \exp\left(\frac{-|\mathbf{v}|^2}{2}\right) d\mathbf{v}}_{= 0} \right. \\ &\quad \left. + \frac{T'}{2} \underbrace{\int_{\mathbb{R}^3} |\mathbf{v}|^2 \exp\left(\frac{-|\mathbf{v}|^2}{2}\right) d\mathbf{v}}_{= 3(2\pi)^{\frac{3}{2}}} \right] \\ &= \rho' ; \end{aligned}$$

$$\begin{aligned}
\bullet \int_{\mathbb{R}^3} v_i \Pi(\rho', \mathbf{u}', T') d\mathbf{v} &= \int_{\mathbb{R}^3} v_i \left[ \rho' - \frac{3}{2} T' + \mathbf{u}' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} \right) T' \right] M[1, \mathbf{0}, 1] d\mathbf{v} \\
&= \frac{1}{(2\pi)^{\frac{3}{2}}} \left[ \left( \rho' - \frac{3}{2} T' \right) \underbrace{\int_{\mathbb{R}^3} v_i \exp \left( \frac{-|\mathbf{v}|^2}{2} \right) d\mathbf{v}}_{=0} + \sum_j u'_j \underbrace{\int_{\mathbb{R}^3} v_i v_j \exp \left( \frac{-|\mathbf{v}|^2}{2} \right) d\mathbf{v}}_{=(2\pi)^{\frac{3}{2}} \delta_{ij}} \right. \\
&\quad \left. + \frac{T'}{2} \underbrace{\int_{\mathbb{R}^3} v_i |\mathbf{v}|^2 \exp \left( \frac{-|\mathbf{v}|^2}{2} \right) d\mathbf{v}}_{=0} \right] \\
&= u'_i ; \\
\bullet \int_{\mathbb{R}^3} |\mathbf{v}|^2 \Pi(\rho', \mathbf{u}', T') d\mathbf{v} &= \int_{\mathbb{R}^3} |\mathbf{v}|^2 \left[ \rho' - \frac{3}{2} T' + \mathbf{u}' \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} \right) T' \right] M[1, \mathbf{0}, 1] d\mathbf{v} \\
&= \frac{1}{(2\pi)^{\frac{3}{2}}} \left[ \left( \rho' - \frac{3}{2} T' \right) \underbrace{\int_{\mathbb{R}^3} |\mathbf{v}|^2 \exp \left( \frac{-|\mathbf{v}|^2}{2} \right) d\mathbf{v}}_{=3(2\pi)^{\frac{3}{2}}} \right. \\
&\quad \left. + \sum_i u'_i \underbrace{\int_{\mathbb{R}^3} v_i |\mathbf{v}|^2 \exp \left( \frac{-|\mathbf{v}|^2}{2} \right) d\mathbf{v}}_{=0} + \frac{T'}{2} \underbrace{\int_{\mathbb{R}^3} |\mathbf{v}|^4 \exp \left( \frac{-|\mathbf{v}|^2}{2} \right) d\mathbf{v}}_{=15(2\pi)^{\frac{3}{2}}} \right] \\
&= 3 \left( \rho' - \frac{3}{2} T' \right) + \frac{15 T'}{2} \\
&= 3 \left( \rho' + T' \right).
\end{aligned}$$

On obtient donc :

$$\left\langle \mathbf{m}(\mathbf{v}) \Pi(\rho', \mathbf{u}', T') \right\rangle = \left( \rho', \mathbf{u}', 3(\rho' + T') \right)^T.$$



## Annexe C

# Modèle à 2 vitesses pour le correcteur de couche limite

L'équation vérifiée par le correcteur s'écrit :

$$v_\eta \frac{\partial \varphi}{\partial \eta} = L(\varphi) = \left[ \rho + \mathbf{u} \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) T \right] M[1, \mathbf{0}, 1] - \varphi, \quad (\text{C.1})$$

avec

$$\rho = \langle \varphi \rangle, \quad \mathbf{u} = \langle \mathbf{v} \varphi \rangle \quad \text{et} \quad T = \left\langle \frac{|\mathbf{v}|^2 - 3}{3} \varphi \right\rangle.$$

On pose :  $\phi = \int_{\mathbb{R}} \varphi dv_\chi$ , ce qui donne

$$\left\{ \begin{array}{l} \rho = \int_{\mathbb{R}^3} \varphi \, dv_\eta dv_\tau dv_\chi = \int_{\mathbb{R}^2} \phi \, dv_\eta \, dv_\tau, \\ u_\eta = \int_{\mathbb{R}^3} v_\eta \varphi \, dv_\eta dv_\tau dv_\chi = \int_{\mathbb{R}^2} v_\eta \phi \, dv_\eta dv_\tau, \\ u_\tau = \int_{\mathbb{R}^3} v_\tau \varphi \, dv_\eta dv_\tau dv_\chi = \int_{\mathbb{R}^2} v_\tau \phi \, dv_\eta dv_\tau, \\ u_\chi = \int_{\mathbb{R}^3} v_\chi \varphi \, dv_\eta dv_\tau dv_\chi = \int_{\mathbb{R}^2} \lambda \, dv_\eta dv_\tau, \\ T = \int_{\mathbb{R}^3} \frac{1}{3} (v_\eta^2 + v_\tau^2 + v_\chi^2 - 3) \varphi \, dv_\eta dv_\tau dv_\chi = \frac{1}{3} \left( \int_{\mathbb{R}^2} (v_\eta^2 + v_\tau^2 - 3) \phi \, dv_\eta dv_\tau + \int_{\mathbb{R}^2} \mu \, dv_\eta dv_\tau \right). \end{array} \right.$$

avec

$$\lambda = \int_{\mathbb{R}} v_\chi \varphi \, dv_\chi, \quad \mu = \int_{\mathbb{R}} v_\chi^2 \varphi \, dv_\chi.$$

On obtient donc trois distributions,  $\phi$ ,  $\lambda$  et  $\mu$ , qui sont les trois premiers moments de  $\varphi$  par rapport à  $v_\chi$ . On écrit une équation d'évolution pour chacune d'entre elles.

**Équation de  $\phi$  :**

On calcule le premier moment de l'équation (C.1) par rapport à  $v_\chi$  :

$$v_\eta \frac{\partial \phi}{\partial \eta} = \rho \int_{\mathbb{R}} M[1, \mathbf{0}, 1] \, dv_\chi + \mathbf{u} \cdot \int_{\mathbb{R}} v M[1, \mathbf{0}, 1] \, dv_\chi + \int_{\mathbb{R}} \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) M[1, \mathbf{0}, 1] \, dv_\chi T - \phi.$$

Or

$$\begin{aligned} \bullet \int_{\mathbb{R}} M[1, \mathbf{0}, 1] \, dv_\chi &= \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbb{R}} \exp\left(-\frac{|\mathbf{v}|^2}{2}\right) \, dv_\chi = \frac{1}{(2\pi)^{\frac{3}{2}}} \exp\left(-\frac{v_\eta^2 + v_\tau^2}{2}\right) \int_{\mathbb{R}} \exp\left(-\frac{v_\chi^2}{2}\right) \, dv_\chi \\ &= \frac{1}{(2\pi)} \exp\left(-\frac{v_\eta^2 + v_\tau^2}{2}\right) \\ &:= \widetilde{M}, \end{aligned}$$

$$\begin{aligned} \bullet \int_{\mathbb{R}} v M[1, \mathbf{0}, 1] \, dv_\chi &= \left( v_\eta \int_{\mathbb{R}} M[1, \mathbf{0}, 1] \, dv_\chi, v_\tau \int_{\mathbb{R}} M[1, \mathbf{0}, 1] \, dv_\chi, \underbrace{\int_{\mathbb{R}} v_\chi M[1, \mathbf{0}, 1] \, dv_\chi}_{=0} \right)^T \\ &= (v_\eta \widetilde{M}, v_\tau \widetilde{M}, 0)^T, \end{aligned}$$

$$\begin{aligned} \bullet \int_{\mathbb{R}} \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) M[1, \mathbf{0}, 1] \, dv_\chi &= \frac{v_\eta^2 + v_\tau^2 - 3}{2} \int_{\mathbb{R}} M[1, \mathbf{0}, 1] \, dv_\chi + \frac{1}{2} \underbrace{\int_{\mathbb{R}} v_\chi^2 M[1, \mathbf{0}, 1] \, dv_\chi}_{= \widetilde{M}} \\ &= \frac{v_\eta^2 + v_\tau^2 - 2}{2} \widetilde{M}. \end{aligned}$$

**Équation de  $\lambda$  :**

On calcule le second moment de l'équation (C.1) par rapport à  $v_\chi$  :

$$v_\eta \frac{\partial \lambda}{\partial \eta} = \rho \int_{\mathbb{R}} v_\chi M[1, \mathbf{0}, 1] \, dv_\chi + \mathbf{u} \cdot \int_{\mathbb{R}} \mathbf{v} v_\chi M[1, \mathbf{0}, 1] \, dv_\chi + \int_{\mathbb{R}} v_\chi \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) M[1, \mathbf{0}, 1] \, dv_\chi T - \lambda.$$

Or

$$\begin{aligned} \bullet \int_{\mathbb{R}} \mathbf{v} v_\chi M[1, \mathbf{0}, 1] \, dv_\chi &= \left( v_\eta \int_{\mathbb{R}} v_\chi M[1, \mathbf{0}, 1] \, dv_\chi, v_\tau \int_{\mathbb{R}} v_\chi M[1, \mathbf{0}, 1] \, dv_\chi, \int_{\mathbb{R}} v_\chi^2 M[1, \mathbf{0}, 1] \, dv_\chi \right)^T \\ &= (0, 0, \widetilde{M})^T, \end{aligned}$$

$$\bullet \int_{\mathbb{R}} v_\chi M[1, \mathbf{0}, 1] \, dv_\chi = \int_{\mathbb{R}} v_\chi \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) M[1, \mathbf{0}, 1] \, dv_\chi = 0 \quad (\text{par imparité}).$$

### Équation de $\mu$ :

Enfin, on obtient l'équation régissant  $\mu$  en calculant le troisième moment de l'équation (C.1) par rapport à  $v_\chi$  :

$$v_\eta \frac{\partial \mu}{\partial \eta} = \rho \int_{\mathbb{R}} v_\chi^2 M[1, \mathbf{0}, 1] dv_\chi + \mathbf{u} \cdot \int_{\mathbb{R}} \mathbf{v} v_\chi^2 M[1, \mathbf{0}, 1] dv_\chi + \int_{\mathbb{R}} v_\chi^2 \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) M[1, \mathbf{0}, 1] dv_\chi T - \mu.$$

Or

$$\bullet \int_{\mathbb{R}} v_\chi^2 M[1, \mathbf{0}, 1] dv_\chi = 1,$$

$$\bullet \int_{\mathbb{R}} \mathbf{v} v_\chi^2 M[1, \mathbf{0}, 1] dv_\chi = \left( v_\eta \int_{\mathbb{R}} v_\chi^2 M[1, \mathbf{0}, 1] dv_\chi, v_\tau \int_{\mathbb{R}} v_\chi^2 M[1, \mathbf{0}, 1] dv_\chi, \int_{\mathbb{R}} v_\chi^3 M[1, \mathbf{0}, 1] dv_\chi \right)^T \\ = (v_\eta \widetilde{M}, v_\tau \widetilde{M}, 0)^T,$$

$$\bullet \int_{\mathbb{R}} v_\chi^2 \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) M[1, \mathbf{0}, 1] dv_\chi = \frac{v_\eta^2 + v_\tau^2 - 3}{2} \int_{\mathbb{R}} v_\chi^2 M[1, \mathbf{0}, 1] dv_\chi + \frac{1}{2} \underbrace{\int_{\mathbb{R}} v_\chi^4 M[1, \mathbf{0}, 1] dv_\chi}_{= 3\widetilde{M}} \\ = \frac{v_\eta^2 + v_\tau^2}{2} \widetilde{M}.$$

### Modèle final

On obtient ainsi le modèle à 3 équations

$$\begin{cases} v_\eta \frac{\partial \phi}{\partial \eta} = \left[ \rho + u_\eta v_\eta + u_\tau v_\tau + \left( \frac{v_\eta^2 + v_\tau^2 - 2}{2} \right) T \right] \widetilde{M} - \phi, \\ v_\eta \frac{\partial \lambda}{\partial \eta} = u_\chi \widetilde{M} - \lambda, \\ v_\eta \frac{\partial \mu}{\partial \eta} = \left[ \rho + u_\eta v_\eta + u_\tau v_\tau + \left( \frac{v_\eta^2 + v_\tau^2}{2} \right) T \right] \widetilde{M} - \mu. \end{cases}$$

On ajoute à ces équations les conditions au bord correspondant à chacun des 3 problèmes aux limites (ref).

- La condition (CL1)  $\varphi|_\Gamma = v_\tau v_\eta M[1, \mathbf{0}, 1]$  donne

$$\begin{cases} \phi|_\Gamma = \int_{\mathbb{R}} v_\tau v_\eta M[1, \mathbf{0}, 1] dv_\chi = v_\tau v_\eta \widetilde{M}, \\ \lambda|_\Gamma = \int_{\mathbb{R}} v_\chi v_\tau v_\eta M[1, \mathbf{0}, 1] dv_\chi = 0, \\ \mu|_\Gamma = \int_{\mathbb{R}} v_\chi^2 v_\tau v_\eta M[1, \mathbf{0}, 1] dv_\chi = v_\tau v_\eta \widetilde{M}, \end{cases}$$



- La condition (CL2)  $\varphi|_{\Gamma} = v_{\tau} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1]$  donne

$$\begin{cases} \phi|_{\Gamma} = \int_{\mathbb{R}} v_{\tau} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1] \, dv_{\chi} = \frac{v_{\tau}}{2} (v_{\eta}^2 + v_{\tau}^2 - 4) \widetilde{M}, \\ \lambda|_{\Gamma} = \int_{\mathbb{R}} v_{\chi} v_{\tau} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1] \, dv_{\chi} = 0, \\ \mu|_{\Gamma} = \int_{\mathbb{R}} v_{\chi}^2 v_{\tau} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1] \, dv_{\chi} = \frac{v_{\tau}}{2} (v_{\eta}^2 + v_{\tau}^2 - 2) \widetilde{M}, \end{cases}$$

- Enfin, la condition (CL3)  $\varphi|_{\Gamma} = v_{\eta} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1]$  donne

$$\begin{cases} \phi|_{\Gamma} = \int_{\mathbb{R}} v_{\eta} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1] \, dv_{\chi} = \frac{v_{\eta}}{2} (v_{\eta}^2 + v_{\tau}^2 - 4) \widetilde{M}, \\ \lambda|_{\Gamma} = \int_{\mathbb{R}} v_{\chi} v_{\eta} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1] \, dv_{\chi} = 0, \\ \mu|_{\Gamma} = \int_{\mathbb{R}} v_{\chi}^2 v_{\eta} \left( \frac{|\mathbf{v}|^2}{2} - \frac{5}{2} \right) M[1, \mathbf{0}, 1] \, dv_{\chi} = \frac{v_{\eta}}{2} (v_{\eta}^2 + v_{\tau}^2 - 2) \widetilde{M}, \end{cases}$$

On dispose également de la condition à l'infini :

$$\frac{\partial \phi}{\partial \eta} \xrightarrow{\eta \rightarrow \infty} 0, \quad \frac{\partial \lambda}{\partial \eta} \xrightarrow{\eta \rightarrow \infty} 0, \quad \frac{\partial \mu}{\partial \eta} \xrightarrow{\eta \rightarrow \infty} 0.$$

On obtient alors les quatre constantes  $K_1, K_2, K_3$  et  $K_4$  en calculant la limite à l'infini des variables du problème. Avec les conditions aux limites (CL1), on a

$$u_{\tau} = \langle v_{\tau} \varphi \rangle \xrightarrow{\eta \rightarrow \infty} K_2 \langle v_{\tau}^2 M[1, \mathbf{0}, 1] \rangle = K_2,$$

de même, avec les conditions (CL2), on obtient

$$u_{\tau} = \xrightarrow{\eta \rightarrow \infty} K_3 \langle v_{\tau}^2 M[1, \mathbf{0}, 1] \rangle = K_3,$$

et enfin, avec les conditions (CL3) :

$$\begin{cases} \rho = \langle \varphi \rangle \xrightarrow{\eta \rightarrow \infty} \left\langle \left( K_1 + K_4 \frac{|\mathbf{v}|^2 - 3}{2} \right) M[1, \mathbf{0}, 1] \right\rangle = K_1, \\ T = \left\langle \frac{|\mathbf{v}|^2 - 3}{3} \varphi \right\rangle \xrightarrow{\eta \rightarrow \infty} \left\langle \frac{|\mathbf{v}|^2 - 3}{3} \left( K_1 + K_4 \frac{|\mathbf{v}|^2 - 3}{2} \right) M[1, \mathbf{0}, 1] \right\rangle = K_4. \end{cases}$$

✎ NB : on peut remarquer que les conditions de bord des 3 problèmes étudiés sont paires par rapport à  $v_\chi$  ; comme l'équation préserve cette symétrie, la solution est également paire en  $v_\chi$ , ce qui implique dans ces cas particuliers  $u_\chi = \langle v_\chi \varphi \rangle = 0$ . Le second membre de l'équation se simplifie donc :

$$L(\varphi) = \left[ \rho + u_\eta v_\eta + u_\tau v_\tau + \left( \frac{|v|^2}{2} - \frac{3}{2} \right) T \right] M[1, \mathbf{0}, 1] - \varphi$$

et la solution  $\lambda \equiv 0$  du modèle à 3 équations est évidente, on peut se contenter de 2 distributions. On voit d'ailleurs que  $\lambda$  ne joue aucun rôle dans l'obtention des constantes.  $\square$



## Annexe D

# Formulaire

On pose  $p(v) = a + \mathbf{b} \cdot \mathbf{v} + \left( \frac{|\mathbf{v}|^2}{2} - \frac{3}{2} \right) c$ . On considère  $(\mathbf{n}, \mathbf{t}, \mathbf{s})$  un repère orthonormal.

- $\left\langle (\mathbf{v} \cdot \mathbf{n}) p(v) M[1, \mathbf{0}, 1] \right\rangle = \mathbf{b} \cdot \mathbf{n}$
- $\left\langle (\mathbf{v} \cdot \mathbf{n})^- p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{-1}{\sqrt{2\pi}} \left( a + \frac{c}{2} \right) + \frac{1}{2} \mathbf{b} \cdot \mathbf{n}$
- $\left\langle (\mathbf{v} \cdot \mathbf{n})^+ p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{1}{\sqrt{2\pi}} \left( a + \frac{c}{2} \right) + \frac{1}{2} \mathbf{b} \cdot \mathbf{n}$
  
- $\left\langle \mathbf{v} (\mathbf{v} \cdot \mathbf{n}) p(v) M[1, \mathbf{0}, 1] \right\rangle = (a + c) \mathbf{n}$
- $\left\langle \mathbf{v} (\mathbf{v} \cdot \mathbf{n})^- p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{1}{2} (a + c) \mathbf{n} - \frac{1}{\sqrt{2\pi}} (\mathbf{n} \otimes \mathbf{n} + \bar{I}) \mathbf{b}$
- $\left\langle \mathbf{v} (\mathbf{v} \cdot \mathbf{n})^+ p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{1}{2} (a + c) \mathbf{n} + \frac{1}{\sqrt{2\pi}} (\mathbf{n} \otimes \mathbf{n} + \bar{I}) \mathbf{b}$
  
- $\left\langle |\mathbf{v}|^2 (\mathbf{v} \cdot \mathbf{n}) p(v) M[1, \mathbf{0}, 1] \right\rangle = 5 \mathbf{b} \cdot \mathbf{n}$
- $\left\langle |\mathbf{v}|^2 (\mathbf{v} \cdot \mathbf{n})^- p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{-4}{\sqrt{2\pi}} (a + c) + \frac{5}{2} \mathbf{b} \cdot \mathbf{n}$
- $\left\langle |\mathbf{v}|^2 (\mathbf{v} \cdot \mathbf{n})^+ p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{4}{\sqrt{2\pi}} (a + c) + \frac{5}{2} \mathbf{b} \cdot \mathbf{n}$
  
- $\left\langle (\mathbf{v} \otimes \mathbf{v}) (\mathbf{v} \cdot \mathbf{n}) p(v) M \right\rangle = (2\mathbf{n} \otimes \mathbf{n} + \bar{I}) \mathbf{b} \cdot \mathbf{n} + (\mathbf{n} \otimes \mathbf{t} + (\mathbf{n} \otimes \mathbf{t})^T) \mathbf{b} \cdot \mathbf{t} + (\mathbf{n} \otimes \mathbf{s} + (\mathbf{n} \otimes \mathbf{s})^T) \mathbf{b} \cdot \mathbf{s}$
- $\left\langle (\mathbf{v} \otimes \mathbf{v}) (\mathbf{v} \cdot \mathbf{n})^- p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{-1}{\sqrt{2\pi}} (\mathbf{n} \otimes \mathbf{n} + \bar{I}) \left( a + \frac{3}{2} c \right) + \frac{1}{2} \left[ (2\mathbf{n} \otimes \mathbf{n} + \bar{I}) \mathbf{b} \cdot \mathbf{n} + (\mathbf{n} \otimes \mathbf{t} + (\mathbf{n} \otimes \mathbf{t})^T) \mathbf{b} \cdot \mathbf{t} + (\mathbf{n} \otimes \mathbf{s} + (\mathbf{n} \otimes \mathbf{s})^T) \mathbf{b} \cdot \mathbf{s} \right]$
- $\left\langle (\mathbf{v} \otimes \mathbf{v}) (\mathbf{v} \cdot \mathbf{n})^+ p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{1}{\sqrt{2\pi}} (\mathbf{n} \otimes \mathbf{n} + \bar{I}) \left( a + \frac{3}{2} c \right) + \frac{1}{2} \left[ (2\mathbf{n} \otimes \mathbf{n} + \bar{I}) \mathbf{b} \cdot \mathbf{n} + (\mathbf{n} \otimes \mathbf{t} + (\mathbf{n} \otimes \mathbf{t})^T) \mathbf{b} \cdot \mathbf{t} + (\mathbf{n} \otimes \mathbf{s} + (\mathbf{n} \otimes \mathbf{s})^T) \mathbf{b} \cdot \mathbf{s} \right]$

- $\left\langle (\mathbf{v} \cdot \mathbf{n}) \mathbf{v} \cdot \nabla_x p(v) M[1, \mathbf{0}, 1] \right\rangle = \nabla (a + c) \cdot \mathbf{n} = \frac{\partial}{\partial n} (a + c)$
- $\left\langle (\mathbf{v} \cdot \mathbf{n})^- \mathbf{v} \cdot \nabla_x p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{1}{2} \frac{\partial}{\partial n} (a + c) - \frac{1}{\sqrt{2\pi}} (\mathbf{n} \otimes \mathbf{n} + \bar{I}) : (\nabla \mathbf{b})^T$   
 $= \frac{1}{2} \frac{\partial}{\partial n} (a + c) - \frac{1}{\sqrt{2\pi}} \left( \frac{\partial}{\partial n} (\mathbf{b} \cdot \mathbf{n}) + \nabla \cdot \mathbf{b} \right)$
- $\left\langle (\mathbf{v} \cdot \mathbf{n})^+ \mathbf{v} \cdot \nabla_x p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{1}{2} \frac{\partial}{\partial n} (a + c) + \frac{1}{\sqrt{2\pi}} (\mathbf{n} \otimes \mathbf{n} + \bar{I}) : (\nabla \mathbf{b})^T$
- $\left\langle \mathbf{v} (\mathbf{v} \cdot \mathbf{n})^- \mathbf{v} \cdot \nabla_x p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{-1}{\sqrt{2\pi}} (\mathbf{n} \otimes \mathbf{n} + \bar{I}) \nabla \left( a + \frac{3}{2} c \right)$   
 $+ \frac{1}{2} \left[ (\nabla \cdot \mathbf{b}) \mathbf{n} + \nabla (\mathbf{b} \cdot \mathbf{n}) + (\nabla \mathbf{b}) \mathbf{n} \right]$   
 $= \frac{-1}{\sqrt{2\pi}} \left[ \nabla \left( a + \frac{3}{2} c \right) + \frac{\partial}{\partial n} \left( a + \frac{3}{2} c \right) \mathbf{n} \right] + \frac{1}{2} \left[ (\nabla \cdot \mathbf{b}) \mathbf{n} + \nabla (\mathbf{b} \cdot \mathbf{n}) + (\nabla \mathbf{b}) \mathbf{n} \right]$
- $\left\langle |\mathbf{v}|^2 (\mathbf{v} \cdot \mathbf{n})^- \mathbf{v} \cdot \nabla_x p(v) M[1, \mathbf{0}, 1] \right\rangle = \frac{\partial}{\partial n} \left( \frac{5}{2} a + 6c \right) - \frac{6}{\sqrt{2\pi}} \left( \frac{\partial}{\partial n} (\mathbf{b} \cdot \mathbf{n}) + \nabla \cdot \mathbf{b} \right)$

# Bibliographie

- [1] A. Alexeenko and S. Chigullapalli. Unsteady 3d rarefied flow solver based on boltzmann-esbgk model kinetic equations. In *41st AIAA Fluid Dynamics Conference*, 2011.
- [2] J. Anderson. *Hypersonic and High-Temperature Gas Dynamics, Second Edition*. AIAA Education Series. American Institute of Aeronautics and Astronautics, 2006.
- [3] P. Andriès, P. Le Tallec, J.-F. Perlat, and B. Perthame. The Gaussian-BGK model of Boltzmann equation with small Prandtl number. *Eur. J. Mech. B/Fluids*, 2000.
- [4] P. Andries, P. Le Tallec, J.-P. Perlat, and B. Perthame. The Gaussian-BGK model of Boltzmann equation with small Prandtl number. *European Journal of Mechanics - B/Fluids*, 19(6) :813 – 830, 2000.
- [5] K. Aoki, P. Degond, and L. Mieussens. Numerical simulations of rarefied gases in curved channels : thermal creep, circulating flow, and pumping effect. *Commun. Comput. Phys.*, 6 :919–954, 2009.
- [6] V. V. Aristov. *Direct methods for solving the Boltzmann equation and study of none-equilibrium flows*, volume 60 of *Fluid Mechanics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.
- [7] V.V. Aristov. Method of adaptative meshes in velocity space for the intense shock wave problem. *USSR J. Comput. Math. Math. Phys.*, 17(4) :261–267, 1977.
- [8] C. Baranger, J. Claudel, N. Hérouard, and L. Mieussens. Deterministic rarefied flow simulation using non cartesian velocity grids. In *ICIAM 2011, Vancouver, Canada*, 2011. Minisymposium "Advanced Numerical Methods for Kinetic Simulations and Their Applications".
- [9] C. Baranger, J. Claudel, N. Hérouard, and L. Mieussens. Locally refined discrete velocity grids for deterministic rarefied flow simulations. *AIP Conference Proceedings*, 1501(1) :389–396, 2012.
- [10] C. Baranger, J. Claudel, N. Hérouard, and L. Mieussens. Locally refined discrete velocity grids for stationary rarefied flow simulations. *Journal of Computational Physics*, 257, Part A(0) :572 – 593, 2014.
- [11] Claude Bardos, Russel E Caflisch, and Basil Nicolaenko. The Milne and Kramers problems for the Boltzmann equation of a hard sphere gas. *Communications on pure and applied mathematics*, 39(3) :323–352, 1986.
- [12] Christophe Berthon and Fabien Marche. A positive preserving high order VFRoe scheme for shallow water equations : A class of relaxation schemes. *SIAM J. Sci. Comput.*, 30(5) :2587–2612, August 2008.

- [13] P.L. Bhatnagar, E.P. Gross, and M. Krook. A model for collision processes in gases. I. small amplitude processes in charged and neutral one-component systems. *Phys. Rev.*, 94 :511–525, 1954.
- [14] G.A. Bird. *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*. Oxford Science Publications, 1994.
- [15] L. Boltzmann. Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Wiener Berichte*, 66 :275 – 370, 1872.
- [16] F. Boyer. Analyse numérique des EDP elliptiques. *Cours de Master 2 Mathématiques et Applications, Université Aix-Marseille*, 2011.
- [17] H. Brezis. *Analyse fonctionnelle, théorie et applications*. Masson-Paris, 1983.
- [18] R. G. Brook. *A parallel, matrix-free newton method for solving approximate boltzmann equations on unstructured topologies*. PhD thesis, The University of Tennessee at Chattanooga, 2008.
- [19] S. Brull and L. Mieussens. Local discrete velocity grids for deterministic rarefied flow simulations. submitted.
- [20] J. M. Burt and I. D. Boyd. A hybrid particle approach for continuum and rarefied flow simulation. *Journal of Computational Physics*, 228 :460–475, 2009.
- [21] C. Cercignani. *Mathematical methods in kinetic theory*. Plenum Press, New York, 1969.
- [22] C. Cercignani and M. Lampis. Kinetic models for gas-surface interactions. *Transport Theory and Statistical Physics*, 1(2) :101–114, 1971.
- [23] S. Chapman and T.G. Cowling. *The Mathematical Theory of Non-Uniform Gases*. Cambridge University Press, 1970.
- [24] S. Chen, K. Xu, C. Lee, and Q. Cai. A unified gas kinetic scheme with moving mesh and velocity space adaptation. *Journal of Computational Physics*, 231(20) :6643 – 6664, 2012.
- [25] C. K. Chu. Kinetic-theoretic description of the formation of a shock wave. *Physics of Fluids (1958-1988)*, 8(1) :12–22, 1965.
- [26] B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab. Local Discontinuous Galerkin methods for the Stokes system. *SIAM Journal on Numerical Analysis*, 40(1) :319 – 343, 2002.
- [27] Bernardo Cockburn and Chi-Wang Shu. Runge–Kutta Discontinuous Galerkin methods for convection-dominated problems. *Journal of Scientific Computing*, 16(3) :173–261, 2001.
- [28] Stéphane Colin. Gas microflows in the slip flow regime : a critical review on convective heat transfer. *Journal of Heat Transfer*, 134(2) :020908, 2012.
- [29] F. Coron. Derivation of slip boundary conditions for the Navier-Stokes system from the Boltzmann equation. *J. Statist. Phys.*, 54(3-4) :829–857, 1989.
- [30] G. Dechristé. *Méthodes numériques sur grilles cartésiennes pour le modèle BGK de l'équation de Boltzmann*. PhD thesis, Université de Bordeaux, 2014. Sous la direction de Luc Mieussens.

- [31] P. Degond and G. Dimarco. Fluid simulations with localized boltzmann upscaling by direct simulation Monte-Carlo. *J. Comput. Phys.*, 231(6) :2414–2437, 2012.
- [32] G. Dimarco and L. Pareschi. Hybrid multiscale methods. II. Kinetic equations. *Multiscale Model. Simul.*, 6(4) :1169–1197, 2008.
- [33] B. Dubroca and L. Mieussens. A conservative and entropic discrete-velocity model for rarefied polyatomic gases. In *CEMRACS 1999 (Orsay)*, volume 10 of *ESAIM Proc.*, pages 127–139 (electronic). Soc. Math. Appl. Indust., Paris, 1999.
- [34] Roland Duclous. *Modélisation et simulation numérique multi-échelle du transport cinétique électronique*. PhD thesis, Université Bordeaux 1, 2009. Sous la direction de Dubroca, Bruno et Tikhonchuk, Vladimir.
- [35] Francis Filbet and Chang Yang. An inverse Lax–Wendroff method for boundary conditions applied to Boltzmann type models. *Journal of Computational Physics*, 245(0) :43 – 61, 2013.
- [36] R. Gatignol. *Théorie Cinétique des Gaz à Répartition Discrète de Vitesses*, volume 36 of *Lecture Notes in Physics*. Springer-Verlag, 1975.
- [37] J.-L. Guermond and G. Kanschat. Asymptotic analysis of upwind Discontinuous Galerkin approximation of the radiative transport equation in the diffusive limit. *SIAM J. Numerical Analysis*, pages 53–78, 2010.
- [38] L. H. Holway. Kinetic theory of shock structure using an ellipsoidal distribution function. In New York Academic Press, editor, *Rarefied Gas Dynamics, Vol. 1 (Proc. Fourth Internat. Sympos. Univ. Toronto, 1964)*, pages 193–215, 1966.
- [39] A. B. Huang and D. L. Hartley. Nonlinear rarefied couette flow with heat transfer. *Phys. Fluids*, 11(6) :1321, 1968.
- [40] J.C. Huang, K. Xu, and P.B. Yu. A unified gas-kinetic scheme for continuum and rarefied flows ii : Multi-dimensional cases. *Communications in Computational Physics*, 3(3) :662–690, 2012.
- [41] J.L.W.V. Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30(1) :175–193, 1906.
- [42] A. Klar. An Asymptotic Preserving numerical scheme for kinetic equations in the low Mach number limit. *SIAM Journal on Numerical Analysis*, 36(5) :1505–1527, 1999.
- [43] M. Knudsen. Die Gesetze der Molekularströmung und der inneren Reibungsströmung der Gase durch Röhren. *Annalen der Physik*, 28 :75–130, 1909.
- [44] M. Knudsen. Eine Revision der Gleichgewichtsbedingung der Gase. Thermische molekulärströmung. *Annalen der Physik*, 31(9) :205–229, 1910.
- [45] V.I. Kolobov and R.R. Arslanbekov. Towards adaptive kinetic-fluid simulations of weakly ionized plasmas. *Journal of Computational Physics*, 231(3) :839 – 869, 2012.
- [46] V.I. Kolobov, R.R. Arslanbekov, V.V. Aristov, A.A. Frolova, and S.A. Zabelok. Unified solver for rarefied and continuum flows with adaptive mesh and algorithm refinement. *Journal of Computational Physics*, 223(2) :589 – 608, 2007.



- [47] V.I. Kolobov, R.R. Arslanbekov, and A.A. Frolova. Boltzmann solver with adaptive mesh in velocity space. In *27th International Symposium on Rarefied Gas Dynamics*, volume 133 of *AIP Conf. Proc.*, pages 928–933. AIP, 2011.
- [48] E. W. Larsen, J.E Morel, and W. F. Miller Jr. Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes. *Journal of Computational Physics*, 69(2) :283 – 324, 1987.
- [49] J. C. Maxwell. Illustrations of the dynamical theory of gases. *Philosophical Magazine Series 4*, 19(124) :19–32, 1860.
- [50] J. C. Maxwell. On stresses in rarified gases arising from inequalities of temperature. *Philosophical Transactions of the Royal Society of London*, 170 :231–256, 1879.
- [51] L. Mieussens. Discrete Velocity Model and Implicit Scheme for the BGK Equation of Rarefied Gas Dynamics. *Math. Models and Meth. in Appl. Sci.*, 8(10) :1121–1149, 2000.
- [52] L. Mieussens. Discrete-velocity models and numerical schemes for the Boltzmann-BGK equation in plane and axisymmetric geometries. *J. Comput. Phys.*, 162 :429–466, 2000.
- [53] Luc Mieussens. *Modèles à vitesses discrètes et méthodes numériques pour l'équation de Boltzmann-BGK*. PhD thesis, Université Bordeaux 1, 1999.
- [54] Luc Mieussens. Discrete-velocity models and numerical schemes for the Boltzmann-BGK equation in plane and axisymmetric geometries. *Journal of Computational Physics*, 162(2) :429 – 466, 2000.
- [55] Luc Mieussens. Convergence of a discrete-velocity model for the Boltzmann-BGK equation. *Computers and Mathematics with Applications*, 41(1–2) :83 – 96, 2001.
- [56] K. Nanbu. Direct simulation scheme derived from the Boltzmann equation. *J. Phys. Soc. of Japan*, 45 :2042–2049, 1980.
- [57] L. Pareschi and G. Russo. Time relaxed Monte Carlo methods for the Boltzmann equation. *SIAM J. Sci. Comput.*, 23(4) :1253–1273 (electronic), 2001.
- [58] B. Perthame. Global existence to the BGK model of Boltzmann Equation. *J. Diff. Eq.*, 82 :191–205, 1989.
- [59] D. A. Di Pietro and A. Ern. *Mathematical aspects of Discontinuous Galerkin methods*, volume 69 of *Mathématiques et Applications*. Springer-Verlag, Berlin Heidelberg, 2012.
- [60] W.H. Reed and T.R. Hill. Triangular mesh methods for the neutron transport equation. *Technical report LA-UR-73-0479*, 1973.
- [61] O. Reynolds. On certain dimensional properties of matter in the gaseous state. Part I and II. *Philosophical Transactions of the Royal Society of London*, 170 :727–845, 1879.
- [62] E.M. Shakhov. Generalization of the Krook kinetic relaxation equation. *Fluid Dynamics*, 3(5) :95–96, 1968.
- [63] Y. Sone. *Kinetic theory and fluid dynamics*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser Boston Inc., Boston, MA, 2002.
- [64] Y. Sone. *Molecular gas dynamics : Theory, techniques, and applications*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser Boston Inc., Boston, MA, 2007.

- [65] H. Struchtrup. The BGK-Model with Velocity-Dependent Collision Frequency. *Continuum Mech. Thermodyn.*, 9 :23–31, 1997.
- [66] H. Struchtrup. Maxwell boundary condition and velocity dependent accommodation coefficient. *Physics of Fluids*, 25(11), 2013.
- [67] V. A. Titarev. Efficient deterministic modelling of three-dimensional rarefied gas flows. *Communications in Computational Physics*, 12(1) :162–192, 2012.
- [68] V.A. Titarev. Conservative numerical methods for model kinetic equations. *Computers Fluids*, 36 :1446–1459, 2007.
- [69] Marcel Vinokur. On one-dimensional stretching functions for finite-difference calculations. *Journal of Computational Physics*, 50(2) :215 – 234, 1983.
- [70] P. Welander. On the temperature jump in a rarefied gas. *Arkiv für Fysik*, 7(44) :507–553, 1954.
- [71] K. Xu and J.-C. Huang. A unified gas-kinetic scheme for continuum and rarefied flows. *J. Comput. Phys.*, 229 :7747–7764, 2010.
- [72] S. Yamamoto and H. Daiguji. Higher order accurate upwind schemes for solving the compressible Euler and Navier-Stokes equations. *Computers Fluids*, 22(2/3) :259–270, 1993.
- [73] H. C. Yee. *A Class of High-Resolution Explicit and Implicit Shock-Capturing Methods*. Von Karman Institute for Fluid Dynamics, Lectures Series, n<sup>o</sup>4. Von Karman Institute for Fluid Dynamics, 1989.